Adv Polym Sci (2005) 185: 1–58
DOI 10.1007/b136794

Published online: 31 August 2005

Incorporating Fluctuations and Dynamics in Self-Consistent Field Theories for Polymer Blends

Marcus Müller[1] (✉) · Friederike Schmid[2]

[1]Department of Physics, University of Wisconsin, 1150 University Avenue, Madison, 53706-1390, USA
mueller@physics.wisc.edu

[2]Fakultät für Physik, Universität Bielefeld, 33615 Bielefeld, Germany
schmid@physik.uni-bielefeld.de

Abstract We review various methods to investigate the statics and the dynamics of collective composition fluctuations in dense polymer mixtures within fluctuating-field

approaches. The central idea of fluctuating-field theories is to rewrite the partition function of the interacting multi-chain systems in terms of integrals over auxiliary, often complex, fields, which are introduced by means of appropriate Hubbard–Stratonovich transformations. Thermodynamic averages such as the average composition and the structure factor can be expressed exactly as averages of these fields. We discuss different analytical and numerical approaches to studying such a theory: The self-consistent field approach solves the integrals over the fluctuating fields in saddle-point approximation. Generalized random phase approximations allow one to incorporate Gaussian fluctuations around the saddle point. Field theoretical polymer simulations are used to study the statistical mechanics of the full system with Complex Langevin or Monte Carlo methods. Unfortunately, they are hampered by the presence of a sign problem. In a dense system, the latter can be avoided without losing essential physics by invoking a saddle point approximation for the complex field that couples to the total density. This leads to the external potential theory. We investigate the conditions under which this approximation is accurate. Finally, we discuss recent approaches to formulate realistic time evolution equations for such models. The methods are illustrated by two examples: A study of the fluctuation-induced formation of a polymeric microemulsion in a polymer-copolymer mixture and a study of early-stage spinodal decomposition in a binary blend.

Keywords Polymer blends · Self-consistent field theory · External potential dynamics · Field-theoretic polymer simulations · Polymeric microemulsion · Polymer dynamics

Abbreviations

α	index for monomer species (A or B)
J	index for polymer species (homopolymer A, homopolymer B, copolymer)
N	number of segments of a chain
χ	Flory–Huggins parameter
$\mathcal{V}$	volume
ϱ	polymer number density
R_e	end-to-end distance of a polymer
$\bar{\mathcal{N}}$	invariant degree of polymerization, $\bar{\mathcal{N}} = \varrho R_\mathrm{e}^d$
$\bar{\phi}_J$	volume fraction of polymers of type J
$\gamma_{J\alpha}$	distributions of segments of species α along a polymer of type J
$\hat{\phi}_\alpha(\boldsymbol{r})$	microscopic segment density of species α (Eq. 3)
$\mathcal{G}$	complex field-theoretical Hamiltonian as a function of fluctuating fields (Eq. 10, 11)
Q_J	single chain partition function of type J (Eqs. 12, 13)
$q_J(\boldsymbol{r},\tau)$	end-segment distribution (Eqs. 17, 18)
$\mathcal{H}$	Hamiltonian of the external potential theory (Eqs. 41, 42)
$\mathcal{F}^*$	Hamiltonian of the (dynamic) SCF theory (Eq. 53)
Λ	Onsager coefficient (Eq. 110)
D	single chain self-diffusion constant

1 Introduction

Polymeric materials in daily life generally are multicomponent systems. Chemically different polymers are "alloyed" as to design a material which combines the favorable characteristics of the individual components [1]. Understanding the miscibility in the molten state is crucial for understanding and tailoring properties that are relevant for practical applications. Miscibility on a microscopic length scale is desirable, for instance, to increase the tensile strength of the material. Unfortunately, different polymers are typically not miscible microscopically, because the entropy of mixing is much smaller for polymers than for small molecules. The mixture separates into domains in which one of the components is enriched. These domains are separated by interfaces. Tuning the interface tension between the coexisting phases, and the morphology of the material on a mesoscopic length scale, is a key to tailoring material properties and has attracted abiding interest.

One strategy to improve the miscibility and the interfacial properties has been to synthesize copolymers by chemically linking polymers of different types to each other. Added in small amounts to a homopolymer mixture, the copolymers adsorb at the interfaces, change the local structure, and reduce the interface tension [2–5]. Mixtures containing substantial amounts of copolymers form new types of phases, where different monomer types aggregate into mesoscopic "phase-separated" domains (microphase separation). By varying the composition of the mixture and the architecture of the individual (co)polymers, one can create a wide range of different morphologies, corresponding to different materials, each with unique material properties.

Unfortunately, copolymers are often rather expensive components. Another, cheaper, strategy for tuning the domain structure of blends takes advantage of the fact that in many practical applications, polymeric blends never reach thermal equilibrium on larger length scales. The morphology of the blend strongly depends on the kinetics of phase separation. Finer dispersed morphologies can be obtained by optimizing the processing conditions [6, 7].

One of the most powerful methods to assess such phenomena theoretically is the self-consistent field (SCF) theory. Originally introduced by Edwards [8] and later Helfand et al. [9], it has evolved into a versatile tool to describe the structure and thermodynamics of spatially inhomogeneous, dense polymer mixtures [10–13]. The SCF theory models a dense multi-component polymer mixture by an incompressible system of Gaussian chains with short-ranged binary interactions and solves the statistical mechanics within the mean-field approximation.

Much of the success of the SCF theory can be traced back to the extended size of the polymer molecules.

First, the large size of the chain molecules imparts a rather universal behavior onto dense polymer mixtures which can be characterized by only a small number of coarse-grained parameters, e.g., the end-to-end distance R_e of the molecules, and the incompatibility per chain χN, where the Flory–Huggins parameter χ [14] describes the repulsion between segments belonging to different components and N is the number of segments per molecule. R_e sets the characteristic length scale of spatial inhomogeneities. Since this length scale is much larger than the size of the monomeric repeat units along the backbone of the polymer chain, these systems can be described successfully by coarse-grained chain models. The parameters, R_e and χN, encode the chemical structure of the polymer on microscopic length scales. Indeed, comparisons between the predictions of the SCF theory and experiments or simulations have shown that the properties of many blends on large length scales depend on the atomistic structure only via the coarse-grained parameters, R_e and χN. The comparison works best if the parameters R_e and χ are considered as "black box" input parameters. Their determination from first principles or even from a well-defined microscopic model (e.g., atomistic force fields), is a formidable theoretical challenge: Even in the simplest case, the Flory–Huggins parameter χ stems from small differences of dispersion forces between the different chemical segments. The value of the end-to-end distance R_e of a polymer in a homogeneous melt results from a subtle screening of excluded volume interactions along the chain by the surrounding molecules, and R_e depends on the density and temperature. Thus χ and R_e cannot be calculated rigorously. If these coarse-grained parameters are determined independently (e.g., by experiments) and used as an input, the SCF theory is successful in making quantitative predictions.

These arguments rationalize the success of a coarse-grained approach. The second reason for the success of the SCF theory, which is a mean-field theory, is the fact that due to their extended shape, the molecules have many interaction partners. Let ϱ denote the polymer number density. The number of molecules inside the volume of a reference chain is then given by $\sqrt{\bar{\mathcal{N}}} = \varrho R_e^d$ (where d denotes the spatial dimension). This quantity measures the degree of interdigitation of the molecules. In a dense three-dimensional melt, $\bar{\mathcal{N}}$ is proportional to the number of segments N per chain, and it is referred to as the invariant degree of polymerization [15]. For systems which differ in $\bar{\mathcal{N}}$ but are characterized by the same χN and R_e the SCF theory will make identical predictions. The quantity $\bar{\mathcal{N}}$ plays an important role as it controls the strength of fluctuations.

In this contribution, we shall discuss improvements which incorporate fluctuations into the SCF theory. Those fluctuations are particularly important (i) in the vicinity of phase transitions (e.g., the unmixing transition in a binary blend or the onset of microphase separation in block copolymers) or (ii) at interfaces, where the translational symmetry is broken and the

local position of the interface can fluctuate. Both types of fluctuations can qualitatively change the behavior compared to the prediction of mean-field theory. The outline of the manuscript is as follows: First we introduce the Edwards model, which is employed as a coarse-grained description of a dense polymer melt throughout the article. Then, we briefly summarize the SCF theory for equilibrium properties and discuss various numerical strategies. Subsequently, we proceed to discuss possibilities to go beyond the mean-field approximation and obtain a description of the dynamics. Finally, we discuss some selected applications and close with a brief outlook on interesting problems.

2
The Edwards Model for Polymer Blends

2.1
Gaussian Model

In dense binary blends, many of the interesting phenomena occur on length scales much larger than the size of a monomeric unit. Hence, a theoretical description can use a coarse-grained model, which only incorporates the relevant ingredients that are necessary to bring about the observed universal behavior: chain connectivity and thermal interaction between unlike monomeric units. The universal properties on length scales much larger than the size of a monomeric unit do not depend on the details of these interactions. One can therefore choose a convenient mathematical model, the Edwards Hamiltonian [8], in which the local microscopic structure enters only via three phenomenological parameters, R_e, χN, and $\bar{\mathcal{N}}$.

In a dense melt, the excluded volume of the monomeric units is screened and chains adopt Gaussian conformations on large length scales. In the following, we shall describe the conformations of a polymer as space curves $\boldsymbol{r}(\tau)$, where the contour parameter τ runs from 0 to 1. The probability distribution $\mathcal{P}[\boldsymbol{r}]$ of such a path $\boldsymbol{r}(\tau)$ is given by the Wiener measure

$$\mathcal{P}[\boldsymbol{r}(\tau)] \sim \exp\left[-\frac{3}{2R_e^2}\int_0^1 \mathrm{d}\tau \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\tau}\right)^2\right]. \tag{1}$$

This describes the Gaussian chain conformations of a non-interacting chain. It is characterized by a single length scale, R_e which is the end-to-end distance of the polymer chain. The structure is Gaussian on *all* length scales, i.e., the model ignores the structure on short length scales [16]: Self-avoiding walk statistics inside the excluded volume blob, rod-like behavior on the length scale of the persistence length, and the details of the monomeric building

units at even shorter length scales as well as long-ranged corrections to the Gaussian behavior due to the screening of excluded volume interactions [17]. All this structure on the microscopic length scale enters the Gaussian chain model only through the single parameter R_e. This parameter (and its dependence on the thermodynamic state, i.e., temperature, density or pressure) cannot be predicted by the theory but is used as input. The Gaussian model provides an appropriate description if the smallest length scale $l_{\min}$ of interest in the calculations is much larger than microscopic length scales. The coarse-graining length scale $l_{\min}$ should be at least on the order of a few persistence lengths. If this condition is violated, the Gaussian chain model can produce qualitatively incorrect predictions (e.g., at low temperatures where both the persistence length becomes large and the interface width becomes small). In this case, other chain models have to be employed that take due account of the local structure (e.g., the worm-like chain model [18], which captures the crossover from Gaussian conformations at large length scales to rod-like behavior on the length scale of the persistence length, or enumeration schemes [19–22] which can deal with arbitrary chain architecture).

In the following we consider a polymer blend of homopolymers of species A and B and/or copolymers containing both types of monomers. For simplicity, we assume that segments of type A and B are perfectly symmetric and that all polymers have the same chain characteristics, in particular, the same chain length. The generalization to more asymmetric situations is straightforward. We characterize the distribution of segments A and B along a chain of given type J (corresponding to a A-homopolymer, a B-homopolymer or an $A:B$ block copolymer) by the segment occupation functions $\gamma_{JA}(\tau)$, $\gamma_{JB}(\tau)$ [23], which take the value 0 or 1 and fulfill

$$\gamma_{JA}(\tau) + \gamma_{JB}(\tau) = 1 \quad \text{for all } J \text{ and } \tau\,. \tag{2}$$

Combining these with the chain conformations $\boldsymbol{r}(\tau)$, we can define a microscopic density $\hat{\phi}_\alpha(\boldsymbol{r})$ for segments of type $\alpha = A$ or B

$$\hat{\phi}_\alpha = \frac{1}{\varrho} \sum_J \sum_{i_J=1}^{n_J} \int_0^1 \mathrm{d}\tau\; \delta(\boldsymbol{r} - \boldsymbol{r}_{i_J}(\tau)) \gamma_{J\alpha}(\tau)\,. \tag{3}$$

Here the sum i_J runs over all n_J polymers of type J, and $\boldsymbol{r}_{i_J}(\tau)$ denotes the conformation of the i_J^{th} polymer. The segment density is normalized by the polymer number density, $\varrho \equiv \sum_J n_J/V$ (V being the volume of the system).

In addition to the chain connectivity, the coarse-grained model has to capture the interactions between segments. In general, a compressible binary polymer mixture [24] exhibits both liquid-liquid immiscibility as well as liquid-vapor type phase separation (cf. *Polymer+solvent systems: Phase diagrams, interface free energies, and nucleation* in this series). In the following, however, we only regard liquid-liquid unmixing into A-rich and B-rich phases

or microphase separation into domains comparable in size to the molecular extension.

To describe the interaction in the polymer liquid, the interaction potential between segments can be decomposed in a short-ranged repulsive part and a longer-ranged attractive part (e.g., using the Barker–Henderson scheme [25] or the Weeks–Chandler–Anderson decomposition [26]). The short-ranged repulsive contribution determines the packing and structure of the fluid. The length scale of fluctuations in the total density is set by the excluded volume correlation length, ξ_{ev}, which also sets the length scale below which the chain conformations do not obey Gaussian but rather self-avoiding walk statistics. In a dense melt, ξ_{ev} is a microscopic length scale. On length scales larger than ξ_{ev} density fluctuations have already decayed, i.e., the fluid is incompressible. For $\xi_{ev} \ll l_{min}$, we can represent the effect of the short-ranged repulsive interactions by an incompressibility constraint:

$$\hat{\phi}_A(\boldsymbol{r}) + \hat{\phi}_B(\boldsymbol{r}) = 1 \tag{4}$$

where we have assumed that both segments occupy identical volumes. Of course, this incompressibility constraint can only be enforced after the coarse-graining procedure, i.e., on length scales larger than ξ_{ev}. The short-ranged repulsive interactions, hence, do not enter the Edwards Hamiltonian explicitly but only via the number density of polymers ϱ. The dependence of ϱ on the thermodynamic state (i.e., the temperature, pressure, and composition of the mixture)—the equation of state—is not predicted by this computational scheme.

The longer-ranged attractive interactions between the segments do not strongly influence the local fluid structure but drive the phase separation. Again we assume that the range of the attractive interactions is small compared to the smallest length scale l_{min} of interest in our calculations. Then, the detailed shape of the longer-ranged potential does not matter either, and we can represent the interactions as zero-ranged. By virtue of the incompressibility constraint, there is only one independent bilinear form of densities and so we use as interaction free energy the expression

$$k_B T \varrho \chi N \int \mathrm{d}^d \boldsymbol{r} \hat{\phi}_A(\boldsymbol{r}) \hat{\phi}_B(\boldsymbol{r}) \tag{5}$$

The combination of the Flory–Huggins parameter and chain length, χN, parameterizes the whole of the subtle interplay between liquid-like structure of the polymeric liquid and the small differences in the attractive interactions between segments of different types. It is worth noting that the Flory–Huggins parameter for neutral polymers without specific interactions is on the order 10^{-4} – 10^{-2} [27], while the strength of the attractive interactions in the fluid are of the order 1 $k_B T$. Hence the χ parameter, which describes the difference in attractive interactions between the species, arises from a strong cancellation. Small changes in the thermodynamic state perturb this balance.

As a result, the χ-parameter depends on the temperature, pressure and composition, as it is often observed in experiments.

Within the coarse-grained model the canonical partition function of a polymer mixture with n_J chains of type J takes the form:

$$Z \sim \frac{1}{\prod_J n_J!} \int \prod_{i=1}^{n} \mathcal{D}[\boldsymbol{r}_i]\mathcal{P}[\boldsymbol{r}_i] \exp\left[-\varrho \int \mathrm{d}^d\boldsymbol{r}\chi N\hat{\phi}_A\hat{\phi}_B\right] \delta(\hat{\phi}_A + \hat{\phi}_B - 1) \quad (6)$$

$$\sim \frac{1}{\prod_J n_J!} \int \prod_J \prod_{i_j=1}^{n_J} \mathcal{D}[\boldsymbol{r}_{i_j}]\mathcal{P}[\boldsymbol{r}_{i_j}] \exp\left[+\frac{\sqrt{\bar{\mathcal{N}}}\chi N}{4R_\mathrm{e}^d} \int \mathrm{d}^d\boldsymbol{r}(\hat{\phi}_A - \hat{\phi}_B)^2\right]$$

$$\times \exp\left[-\frac{\sqrt{\bar{\mathcal{N}}}\chi N}{4R_\mathrm{e}^d} \int \mathrm{d}^d\boldsymbol{r}(\hat{\phi}_A + \hat{\phi}_B)^2\right] \delta(\hat{\phi}_A + \hat{\phi}_B - 1)$$

The product $\prod_i$ runs over all $n = \sum n_J$ chains, and the functional integral $\mathcal{D}[\boldsymbol{r}_i]$ sums over all possible conformations of the chain i. Within the Edwards Hamiltonian, the system is thus described by the following parameters: χN describes the incompatibility of the different polymer species, R_e sets the size of the molecules, $\bar{\mathcal{N}}$ is the invariant chain length, which describes the degree of interdigitation of the molecules, $\bar{\phi}_J \equiv n_J/\varrho V$ is the average density of type J, and $\gamma_{J\alpha}(\tau)$ the distribution of segments along the molecules of type J. The quantity $\bar{\mathcal{N}}$ sets the scale of the free energy but does not influence the SCF solution (i.e., the location of the extrema of the free energy, see Sect. 3).

2.2 Hubbard–Stratonovich Transformation and Fluctuating Fields

The partition function Eq. 6 describes a system of mutually interacting chains. Introducing auxiliary fields, U and W, via a Hubbard–Stratonovich transformation, one can decouple the interaction between the chains and rewrite the Hamiltonian in terms of independent chains in fluctuating fields. Then, one can integrate over the chain conformations and obtain a Hamiltonian which only depends on the auxiliary fields. Thermodynamic averages like density or structure factors can be expressed as averages over the fields, U and W, without approximation.

In the following we introduce a minimal transformation to decouple the interactions between polymer chains. Two different schemes have to be employed for the thermal interactions between the monomers and the incompressibility constraint. For the thermal interactions which give rise to the term $(\hat{\phi}_A - \hat{\phi}_B)^2$ in the Hamiltonian, we use the Hubbard–Stratonovich formula

$$\exp[\alpha x^2] = \frac{1}{2\sqrt{\pi\alpha}} \int_{-\infty}^{+\infty} \mathrm{d}y \exp\left[-\left(\frac{y^2}{4\alpha} + xy\right)\right] \quad (7)$$

at each point in space and identify $x = \hat{\phi}_A - \hat{\phi}_B$, $y = \varrho W/2$, and $\alpha = \sqrt{\bar{N}}\chi N/4R_e^d$. This introduces a real auxiliary field W.

To rewrite the incompressibility constraint, we use the Fourier representation of the δ-function:

$$\delta(x-1) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}y \exp\left[-\mathrm{i}y(x-1)\right] \tag{8}$$

using $x = \hat{\phi}_A + \hat{\phi}_B$ and $y = \varrho U/2$. This introduces a real auxiliary field U.

The field W, which couples to the composition of the mixture, gives rise to a real contribution in the exponent, while the field U, which enforces incompressibility, yields a complex contribution.

We note that this complex term cannot be avoided easily. Replacing the incompressibility constraint by a finite compressibility

$$\varrho N\frac{\kappa}{2}\int \mathrm{d}^d\boldsymbol{r}(\hat{\phi}_A + \hat{\phi}_A - 1)^2 \tag{9}$$

where κ denotes the isothermal compressibility of the polymer liquid, and performing the appropriate Hubbard–Stratonovich transformation would also introduce a complex term in the exponential (cf. Sect. 4.4.2).

Using these two expressions we can rewrite the partition function exactly as

$$\begin{aligned} Z &\sim \frac{1}{\prod_J n_J!}\int \prod_i \mathcal{D}[\boldsymbol{r}_i]\mathcal{P}[\boldsymbol{r}_i]\mathcal{D}U[\boldsymbol{r}]\mathcal{D}W[\boldsymbol{r}] \exp\left[-\frac{\varrho V\chi N}{4}\right] \\ &\quad\times \exp\left[-\varrho\int \mathrm{d}^d\boldsymbol{r}\left\{\frac{W^2}{4\chi N} + \frac{W}{2}(\hat{\phi}_A - \hat{\phi}_B) + \frac{iU}{2}\left(\hat{\phi}_A + \hat{\phi}_B - 1\right)\right\}\right] \\ &\sim \int \mathcal{D}U\mathcal{D}W \exp\left[-\frac{\mathcal{G}[U,W]}{k_\mathrm{B}T}\right], \end{aligned} \tag{10}$$

which defines a Hamiltonian $\mathcal{G}[U, W]$ [28, 29] for the fields U and W,

$$\frac{\mathcal{G}[U,W]}{\sqrt{\bar{N}}k_\mathrm{B}T(V/R_e^d)} = -\sum_J \bar{\phi}_J \ln\frac{VQ[W_A,W_B]}{n_J} + \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\frac{W^2}{4\chi N}. \tag{11}$$

Here $Q_J[W_A, W_B]$ denotes the partition function of a single non-interacting Gaussian chain of type J

$$Q_J[W_A,W_B] = \frac{1}{V}\int \mathcal{D}[\boldsymbol{r}]\mathcal{P}[\boldsymbol{r}] \exp\left[-\int_0^1 \mathrm{d}\tau \sum_{\alpha=A,B} W_\alpha(\boldsymbol{r}(\tau))\gamma_{J\alpha}(\tau)\right] \tag{12}$$

$$= \frac{1}{V}\int \mathcal{D}[\boldsymbol{r}]\mathcal{P}[\boldsymbol{r}] \exp\left[-\frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\sum_\alpha W_\alpha(\boldsymbol{r})\hat{\phi}_{J\alpha}^{\mathrm{sc}}(\boldsymbol{r})\right] \tag{13}$$

in the fields

$$W_A = \frac{\mathrm{i}U + W}{2} \qquad W_B = \frac{\mathrm{i}U - W}{2} \,. \tag{14}$$

Here we have normalized Q_J by the volume V in order to make the trivial contribution of the translational entropy explicit. The dimensionless, microscopic single chain density is defined as

$$\hat{\phi}^{\mathrm{sc}}_{J\alpha}(\boldsymbol{r}) = V \int \mathrm{d}\tau \delta(\boldsymbol{r} - \boldsymbol{r}(\tau)) \tag{15}$$

where $\alpha = A$ or B describes the segment species and J denotes the type of polymer. In Eq. 11, we have omitted a term $-(1/2V)\int \mathrm{d}^d\boldsymbol{r}\left(\mathrm{i}U - \chi N/2\right)$. This is legitimated by the fact that $\mathcal{G}[U, W]$ is invariant under adding a spatially constant (real or imaginary) field $\Delta\bar{U}$ to U. Therefore, we can fix

$$\frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\mathrm{i}U(\boldsymbol{r}) = \frac{\chi N}{2} \,. \tag{16}$$

To calculate the single chain partition function Q_J, it is useful to define the end segment distribution, $q_J(\boldsymbol{r}, t)$, which describes the probability of finding the end of a chain fragment—containing the segments $[0:t]$ and being exposed to the fields $W_\alpha(\alpha = A, B)$ —at position $\boldsymbol{r}$ [9]:

$$q_J(\boldsymbol{r}, t) = \int_0^t \mathcal{D}[\boldsymbol{r}]\mathcal{P}[\boldsymbol{r}]\delta(\boldsymbol{r}(t) - \boldsymbol{r})\, \mathrm{e}^{-\int_0^t \mathrm{d}\tau \sum_\alpha W_\alpha(\boldsymbol{r}(\tau))\gamma_{J\alpha}(\tau)} \tag{17}$$

Similarly, we define

$$q^\dagger_J(\boldsymbol{r}, t) = \int_t^1 \mathcal{D}[\boldsymbol{r}]\mathcal{P}[\boldsymbol{r}]\delta(\boldsymbol{r}(t) - \boldsymbol{r})\, \mathrm{e}^{-\int_t^1 \mathrm{d}\tau \sum_\alpha W_\alpha(\boldsymbol{r}(\tau))\gamma_{J\alpha}(\tau)} \,. \tag{18}$$

The end segment distributions obey the following diffusion equation [9]

$$\frac{\partial q_J(\boldsymbol{r}, t)}{\partial t} = \frac{R_\mathrm{e}^2}{6}\nabla^2 q_J(\boldsymbol{r}, t) - \sum_{\alpha=A,B} W_\alpha(\boldsymbol{r})\gamma_{J\alpha}(t) q_J(\boldsymbol{r}, t) \tag{19}$$

$$-\frac{\partial q^\dagger_J(\boldsymbol{r}, t)}{\partial t} = \frac{R_\mathrm{e}^2}{6}\nabla^2 q^\dagger_J(\boldsymbol{r}, t) - \sum_{\alpha=A,B} W_\alpha(\boldsymbol{r})\gamma_{J\alpha}(t) q^\dagger_J(\boldsymbol{r}, t) \tag{20}$$

with the boundary condition $q_J(\boldsymbol{r}, 0) = 1$ and $q^\dagger_J(\boldsymbol{r}, 1) = 1$, i.e., the beginning of the chain fragment is uniformly distributed. For homopolymers the two propagators are related via $q_J(\boldsymbol{r}, t) = q^\dagger_J(\boldsymbol{r}, 1 - t)$. The solution of these equations yields the single chain partition function:

$$Q_J = \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r} q_J(\boldsymbol{r}, 1) = \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r} q^\dagger_J(\boldsymbol{r}, 0) = \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r} q_J(\boldsymbol{r}, t) q^\dagger_J(\boldsymbol{r}, t) \quad \forall t \tag{21}$$

We note that the fluctuations in the spatially homogeneous component $\bar{W} \equiv (1/V)\int \mathrm{d}^d\boldsymbol{r} W$ of the field W are strictly Gaussian. To show this, we decompose $W = \bar{W} + W'$ and obtain:

$$\frac{\mathcal{G}[U,W]}{\sqrt{\mathcal{N}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} = \frac{\mathcal{G}[U,W']}{\sqrt{\mathcal{N}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} + \frac{1}{4\chi N}\left[\bar{W} + \chi N(\bar{\phi}_A - \bar{\phi}_B)\right]^2 - \frac{\chi N}{4}(\bar{\phi}_A - \bar{\phi}_B)^2 \tag{22}$$

The only effect of a spatially homogeneous field acting on A-segments or B-segments is to introduce additional contributions to the chemical potentials μ_J—quantities which are immaterial in the canonical ensemble.

By the Hubbard–Stratonovich transformation we have rewritten the partition function of the interacting multi-chain systems in terms of non-interacting chains in complex fluctuating fields, $\mathrm{i}U + W$ and $\mathrm{i}U - W$. In field theoretical polymer simulations, one samples the fields U and W via computer simulation using the above Hamiltonian (cf. Sect. 4.4).

To calculate thermal averages of $\hat{\phi}_A(\boldsymbol{r}) - \hat{\phi}_B(\boldsymbol{r})$ we introduce a local exchange potential, $\Delta\mu$, that couples to the local composition, $\hat{\phi}_A - \hat{\phi}_B$ [29]:

$$\begin{aligned}\widetilde{Z}[\Delta\mu] &\sim \frac{1}{\prod_J n_J!}\int\prod_i \mathcal{D}[\boldsymbol{r}_i]\mathcal{P}[\boldsymbol{r}_i]\mathcal{D}U[\boldsymbol{r}]\mathcal{D}W[\boldsymbol{r}] \\ &\quad\times\exp\left[-\varrho\int\mathrm{d}^d\boldsymbol{r}\left\{\frac{W^2}{4\chi N} + \frac{W+\Delta\mu}{2}(\hat{\phi}_A - \hat{\phi}_B) + \frac{\mathrm{i}U}{2}\left(\hat{\phi}_A + \hat{\phi}_B\right)\right\}\right] \\ &\sim \int\mathcal{D}U\mathcal{D}W\exp\left[-\frac{\tilde{\mathcal{G}}[U,W,\Delta\mu]}{k_\mathrm{B}T}\right]\end{aligned} \tag{23}$$

The free energy functional $\tilde{\mathcal{G}}[U,W,\Delta\mu]$ takes the form

$$\begin{aligned}\frac{\tilde{\mathcal{G}}[U,W,\Delta\mu]}{\sqrt{\mathcal{N}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} &= -\sum_J\bar{\phi}_J\ln\frac{VQ_J[\frac{1}{2}(\mathrm{i}U + W + \Delta\mu), \frac{1}{2}(\mathrm{i}U - W - \Delta\mu)]}{n_J} \\ &\quad + \frac{1}{V}\int\mathrm{d}^d\boldsymbol{r}\frac{W^2}{4\chi N} \\ &= \frac{\mathcal{G}[U,W]}{\sqrt{\mathcal{N}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} + \frac{1}{4\chi NV}\int\mathrm{d}^d\boldsymbol{r}(\Delta\mu^2 - 2\Delta\mu W)\end{aligned} \tag{24}$$

where we have changed the dummy field W to $W + \Delta\mu$ in the last step. Differentiating Eq. 24 with respect to $\Delta\mu(\boldsymbol{r})$ we obtain:

$$\langle\hat{\phi}_A(\boldsymbol{r}) - \hat{\phi}_B(\boldsymbol{r})\rangle = -\frac{2}{\varrho}\frac{1}{\widetilde{Z}[\Delta\mu]}\frac{\delta\widetilde{Z}[\Delta\mu]}{\delta\Delta\mu(\boldsymbol{r})}\bigg|_{\Delta\mu=0} = -\frac{1}{\chi N}\langle W\rangle \tag{25}$$

This expression relates the thermodynamic average of the composition difference to the average of the real field W. Similarly, one can generate higher

moments of the composition:

$$\langle[\hat{\phi}_A(\boldsymbol{r})-\hat{\phi}_B(\boldsymbol{r})][\hat{\phi}_A(\boldsymbol{r}')-\hat{\phi}_B(\boldsymbol{r}')]\rangle = \frac{4}{\varrho^2}\frac{1}{\widetilde{Z}[\Delta\mu]}\left.\frac{\delta^2\widetilde{Z}[\Delta\mu]}{\delta\Delta\mu(\boldsymbol{r})\delta\Delta\mu(\boldsymbol{r}')}\right|_{\Delta\mu=0} \tag{26}$$
$$= -\frac{2R_e^d\delta(\boldsymbol{r}-\boldsymbol{r}')}{\sqrt{\mathcal{N}}\chi N}+\frac{\langle W(\mathbf{r})W(\mathbf{r}')\rangle}{(\chi N)^2}$$

The two expressions Eqs. 25 and 26 allow us to calculate the physically important thermodynamic average of the microscopic densities and their fluctuations from the thermodynamics average and fluctuations of the field W. Although the Hamiltonian $\mathcal{G}[U,W]$ is complex, the thermodynamic average of the microscopic densities and their fluctuations can be expressed in terms of the real field W [29].

Alternatively, one can calculate the average of the microscopic densities from Eq. 24 and obtain

$$\langle\hat{\phi}_A(\mathbf{r})-\hat{\phi}_B(\mathbf{r})\rangle = \langle\phi_A^*[W_A,W_B]-\phi_B^*[W_A,W_B]\rangle\,, \tag{27}$$

where the ϕ_α^* ($\alpha=A,B$) are functionals of $W_A=(\mathrm{i}U+W)/2$ and $W_B=(\mathrm{i}U-W)/2$ (cf. Eq. 14) given by

$$\phi_\alpha^*(\boldsymbol{r}) = -\frac{1}{\varrho}\sum_J n_J\frac{\delta\ln VQ_J[W_A,W_B]}{\delta W_\alpha(\boldsymbol{r})} \tag{28}$$
$$= \sum_J \bar{\phi}_J\frac{\int\mathcal{D}[\boldsymbol{r}]\mathcal{P}[\boldsymbol{r}]\,\mathrm{e}^{-\frac{1}{V}\int\mathrm{d}^d\boldsymbol{r}\sum_\alpha W_\alpha(\boldsymbol{r})\hat{\phi}_{J\alpha}^{\mathrm{sc}}(\boldsymbol{r})}V\int\mathrm{d}\tau\delta(\boldsymbol{r}-\boldsymbol{r}(\tau))\gamma_{J\alpha}(\tau))}{\int\mathcal{D}[\boldsymbol{r}]\mathcal{P}[\boldsymbol{r}]\,\mathrm{e}^{-\frac{1}{V}\int\mathrm{d}^d\boldsymbol{r}\sum_\alpha W_\alpha(\boldsymbol{r})\hat{\phi}_{J\alpha}^{\mathrm{sc}}(\boldsymbol{r})}}$$
$$\equiv \sum_J \bar{\phi}_J\langle\hat{\phi}_{J\alpha}^{\mathrm{sc}}(\boldsymbol{r})\rangle\,. \tag{29}$$

This equation identifies ϕ_α^* as the density of α-segments created by a single chain in the external fields W_A and W_B, averaged over all polymer types J. (We note that the average normalized microscopic density of a single chain in $\langle\hat{\phi}^{\mathrm{sc}}\rangle$ in a volume $\mathcal{V}$ is independent from $\mathcal{V}$, cf. Eq. 14.) The thermodynamic average of the composition is simply the Boltzmann average of the corresponding single chain properties averaged over the fluctuating fields. The functional derivatives $\phi_A^*(\boldsymbol{r})$ and $\phi_B^*(\boldsymbol{r})$ can be calculated using the end-segment distribution functions, $q_J(\boldsymbol{r},t)$ and $q_J^\dagger(\boldsymbol{r},t)$, according to:

$$\phi_\alpha^*(\boldsymbol{r}) = \sum_J\bar{\phi}_J\frac{1}{Q_J}\int_0^1\mathrm{d}t q_J(\boldsymbol{r},t)q_J^\dagger(\boldsymbol{r},t)\gamma_{J\alpha}(t) \tag{30}$$

While the thermodynamic average of the microscopic composition $\hat{\phi}_\alpha$ equals the average of the functional ϕ_α^* over the fluctuating fields W_A and W_B, or, alternatively, U and W, such a simple relation does not hold true for the fluc-

tuations. Taking the second derivative with respect to $\Delta\mu$, we obtain:

$$\langle[\hat{\phi}_A(\boldsymbol{r})-\hat{\phi}_B(\boldsymbol{r})][\hat{\phi}_A(\boldsymbol{r}')-\hat{\phi}_B(\boldsymbol{r}')]\rangle = \langle[\phi_A^*(\boldsymbol{r})-\phi_B^*(\boldsymbol{r})][\phi_A^*(\boldsymbol{r}')-\phi_B^*(\boldsymbol{r}')]\rangle - \frac{R_e^d}{\sqrt{\mathcal{N}}}\left\langle \frac{\delta\phi_A^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')}+\frac{\delta\phi_B^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')} - \frac{\delta\phi_A^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')}-\frac{\delta\phi_B^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')}\right\rangle \quad (31)$$

The fluctuations of the physically relevant microscopic monomer densities, $\hat{\phi}_A$ and $\hat{\phi}_B$, and the fluctuations of the Boltzmann averaged single chain properties, ϕ_A^* and ϕ_B^*, due to fluctuations of the fields U and W are not identical. The additional term accounts for the single chain correlations [30]:

$$R_e^d\frac{\delta\phi_\alpha^*(\boldsymbol{r})}{\delta W_\beta(\boldsymbol{r}')} = \sum_J \frac{\bar{\phi}_J R_e^d}{V}\left(\langle\hat{\phi}_{J\alpha}^{sc}(\boldsymbol{r})\hat{\phi}_{J\beta}^{sc}(\boldsymbol{r}')\rangle - \langle\hat{\phi}_{J\alpha}^{sc}(\boldsymbol{r})\rangle\langle\hat{\phi}_{J\beta}^{sc}(\boldsymbol{r}')\rangle\right). \quad (32)$$

We emphasize that the additional term due to single chain correlations is of the same order of magnitude as the fluctuations of the physically relevant microscopic monomer densities. We shall demonstrate this explicitly in the framework of the Random-Phase Approximation (RPA) in Sect. 4.2.

To calculate thermal averages of $\hat{\phi}_A(\boldsymbol{r})+\hat{\phi}_B(\boldsymbol{r})$ we introduce a spatially varying total chemical potential $\delta\mu$

$$\begin{aligned}\widetilde{Z}[\delta\mu] &\sim \frac{1}{\prod_J n_J!}\int\prod_i \mathcal{D}[\boldsymbol{r}_i]\mathcal{P}[\boldsymbol{r}_i]\mathcal{D}U[\boldsymbol{r}]\mathcal{D}W[\boldsymbol{r}] \quad (33)\\ &\quad\times\exp\left[-\varrho\int \mathrm{d}^d\boldsymbol{r}\left\{\frac{W^2}{4\chi N}+\frac{W}{2}(\hat{\phi}_A-\hat{\phi}_B)+\frac{\mathrm{i}U+\delta\mu}{2}\left(\hat{\phi}_A+\hat{\phi}_B\right)\right\}\right]\\ &\sim\int\mathcal{D}U\mathcal{D}W\exp\left[-\frac{\tilde{\mathcal{G}}[U,W,\delta\mu]}{k_BT}\right]\end{aligned}$$

Similar to Eqs.24 and 25 we can rewrite the free energy functional $\tilde{\mathcal{G}}[U,W,\delta\mu]$ as:

$$\frac{\tilde{\mathcal{G}}[U,W,\delta\mu]}{\sqrt{\mathcal{N}}k_BT(V/R_e^d)} = \frac{\mathcal{G}[U,W]}{\sqrt{\mathcal{N}}k_BT(V/R_e^d)}+\frac{1}{2V}\int\mathrm{d}^d\boldsymbol{r}\delta\mu \quad (34)$$

$$= -\sum_J\bar{\phi}_J\ln\frac{VQ_J}{n_J}+\frac{1}{V}\int\mathrm{d}^d\boldsymbol{r}\frac{W^2}{4\chi N} \quad (35)$$

The first derivative *wrt* $\delta\mu$ yields the average of the local monomer density

$$\langle\hat{\phi}_A(\boldsymbol{r})+\hat{\phi}_B(\boldsymbol{r})\rangle = -\frac{2}{\varrho}\frac{1}{\widetilde{Z}[\delta\mu]}\frac{\delta\widetilde{Z}[\delta\mu]}{\delta\delta\mu(\boldsymbol{r})}\bigg|_{\delta\mu=0} = \langle\phi_A^*+\phi_B^*\rangle \quad (36)$$

$$= 1 \quad (37)$$

The second derivative yields the density fluctuations:

$$\langle[\hat{\phi}_A(\boldsymbol{r}) + \hat{\phi}_B(\boldsymbol{r})][\hat{\phi}_A(\boldsymbol{r}') + \hat{\phi}_B(\boldsymbol{r}')]\rangle = \frac{4}{\varrho^2}\frac{1}{\widetilde{Z}[\delta\mu]}\left.\frac{\delta^2\widetilde{Z}[\delta\mu]}{\delta\delta\mu(\boldsymbol{r})\delta\delta\mu(\boldsymbol{r}')}\right|_{\delta\mu=0}$$

$$= \langle[\phi_A^*(\boldsymbol{r}) + \phi_B^*(\boldsymbol{r})][\phi_A^*(\boldsymbol{r}') + \phi_B^*(\boldsymbol{r}')]\rangle$$

$$- \frac{R_e^d}{\sqrt{\mathcal{N}}}\left\langle\frac{\delta\phi_A^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')} + \frac{\delta\phi_B^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')} + \frac{\delta\phi_A^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')} + \frac{\delta\phi_B^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')}\right\rangle \tag{38}$$

$$= 1 \tag{39}$$

Note that the last contribution in the equation above is similar to the single chain correlations in Eq. 31. The incompressibility constraint is enforced on the microscopic density $\hat{\phi}_A + \hat{\phi}_B$. At this stage, ϕ_A^* and ϕ_B^* are only auxiliary functionals of the fields U, W, which are proportional to the density distribution of a single chain in the corresponding external fields.

3 External Potential (EP) Theory and Self-Consistent Field (SCF) Theory

The reformulation of the partition function in terms of single chains in the fluctuating, complex fields $W_A = (\mathrm{i}U + W)/2$ and $W_B = (\mathrm{i}U - W)/2$ is exact. The numerical evaluation of the functional integral over the fields U and W is, however, difficult (cf. Sect. 4.4), and various approximations have been devised.

The functional integral over the fluctuating field U, conjugated to the total density, can be approximated by the "saddle point": The integrand evaluated at that function $U^*[W]$ which minimizes the free energy function $\mathcal{G}[U, W]$. Carrying out this saddle point integration in the field U, we obtain a free energy functional $\mathcal{H}[W]$. In the following we denote this approach external potential (EP) theory following Maurits and Fraaije who derived the saddle point equations heuristically [31]. This scheme still retains the important fluctuations in the field W, conjugated to the composition. Using an additional saddle point approximation for W, we neglect fluctuations in the composition and we arrive at the self-consistent field theory.

3.1 External Potential (EP) Theory: Saddle Point Integration in U

The fluctuations described by the two fields, U and W, are qualitatively different. This is already apparent from the fact that one, U, gives rise to a complex contribution to the field that acts on a chain, while the other, W, corresponds to a real one. The field U couples to the total density $\hat{\phi}_A + \hat{\phi}_B$ and has been introduced to decouple the incompressibility constraint. Qualitatively, it con-

trols the fluctuations of the total density, and it does not directly influence the expectation value or the fluctuations of the composition $\hat{\phi}_A - \hat{\phi}_B$.

The saddle point value U^* is given by the condition:

$$\frac{\delta G[U^*, W]}{\delta U} = 0 \quad \Rightarrow \quad \phi_A^* + \phi_B^* = 1 \,. \tag{40}$$

Instead of enforcing the incompressibility constraint on each of the microscopic conformations, we thus only require that the single chain averages in the external field obey the constraint. We recall that the ϕ_α^* are functionals of $W_A = (\mathrm{i}U + W)/2$ and $W_B = (\mathrm{i}U - W)/2$ (cf. Eq. 28), thus Eq. 40 implicitly defines a functional $U^*[W]$. Substituting the saddle point value into the free energy functional (Eq. 11), we obtain an approximate partition function

$$Z_{\mathrm{EP}} \sim \int \mathcal{D}W \exp\left[-\frac{\mathcal{H}[W]}{k_{\mathrm{B}}T}\right] \tag{41}$$

with the free energy functional

$$\begin{aligned} \frac{\mathcal{H}[W]}{\sqrt{\bar{\mathcal{N}}}k_{\mathrm{B}}T(V/R_{\mathrm{e}}^d)} &= \frac{\mathcal{G}[U^*, W]}{\sqrt{\bar{\mathcal{N}}}k_{\mathrm{B}}T(V/R_{\mathrm{e}}^d)} \\ &= -\sum_J \bar{\phi}_J \ln\frac{VQ_J}{n_J} + \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\frac{W^2}{4\chi N} \end{aligned} \tag{42}$$

At this point, one has two possible choices for calculating the thermodynamic averages of the microscopic composition [29]:

1. Although one performs a saddle point approximation in U, one can use the expressions Eq. 25 and Eq. 26, which relate the average of the field W and its fluctuations to the average and the fluctuations of the microscopic composition. The crucial difference is, however, that the fluctuations of the field W are now governed by the Hamiltonian $\mathcal{H}[W]$ of the EP theory and not by the exact Hamiltonian $\mathcal{G}[U, W]$. If the saddle point approximation is reliable, $\mathcal{G}[U, W]$ will be well described by a parabola in $U - U^*[W]$, the fluctuations in U will be nearly Gaussian and will hardly influence the fluctuations of W. Therefore, the fluctuations of W generated by the Hamiltonian $\mathcal{H}[W]$ of the EP theory closely mimic the fluctuations of the field W corresponding to the exact Hamiltonian $\mathcal{G}[U, W]$, i.e., $\langle W(\boldsymbol{r})W(\boldsymbol{r}')\rangle_{\mathrm{EP}} \approx \langle W(\boldsymbol{r})W(\boldsymbol{r}')\rangle$. By the same token, using Eq. 26, we ensure that the composition fluctuations in the EP theory closely resemble the exact thermodynamic average of fluctuations of the microscopic composition.

2. One can introduce a spatially varying exchange potential $\delta\mu$ and calculate the thermodynamic averages with respect to the free energy functional $\mathcal{H}[W]$ of the EP theory. This leads to expressions similar to Eq. 27 and

Eq. 31:

$$\langle \hat{\phi}_A(\boldsymbol{r}) - \hat{\phi}_B(\boldsymbol{r}) \rangle_{\mathrm{EP}} = \langle \phi_A^* - \phi_B^* \rangle_{\mathrm{EP}} \tag{43}$$

$$\begin{aligned} &\langle [\hat{\phi}_A(\boldsymbol{r}) - \hat{\phi}_B(\boldsymbol{r})][\hat{\phi}_A(\boldsymbol{r}') - \hat{\phi}_B(\boldsymbol{r}')] \rangle_{\mathrm{EP}} \\ &\quad = \langle \phi_A^*(\boldsymbol{r}) - \phi_B^*(\boldsymbol{r})][\phi_A^*(\boldsymbol{r}') - \phi_B^*(\boldsymbol{r}')] \rangle_{\mathrm{EP}} \\ &\quad - \frac{R_{\mathrm{e}}^d}{\sqrt{\mathcal{N}}} \left\langle \frac{\delta \phi_A^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')} + \frac{\delta \phi_B^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')} - \frac{\delta \phi_A^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')} - \frac{\delta \phi_B^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')} \right\rangle_{\mathrm{EP}} \end{aligned} \tag{44}$$

We refer to the last expression as the literal composition fluctuations in the EP theory. Likewise, one can use the functional derivative with respect to a spatially varying chemical potential $\delta\mu$, which couples to $\hat{\phi}_A + \hat{\phi}_B$, to calculate the literal average of the total density.

$$\langle \hat{\phi}_A(\boldsymbol{r}) + \hat{\phi}_B(\boldsymbol{r}) \rangle_{\mathrm{EP}} = \langle \phi_A^* + \phi_B^* \rangle_{\mathrm{EP}} = 1 \tag{45}$$

For the literal fluctuations of the total density in the EP theory (cf. Eq. 38) one obtains:

$$\begin{aligned} &\langle [\hat{\phi}_A(\boldsymbol{r}) + \hat{\phi}_B(\boldsymbol{r})][\hat{\phi}_A(\boldsymbol{r}') + \hat{\phi}_B(\boldsymbol{r}')] \rangle_{\mathrm{EP}} \\ &\quad = \langle \underbrace{[\phi_A^*(\boldsymbol{r}) + \phi_B^*(\boldsymbol{r})]}_{=1} \underbrace{[\phi_A^*(\boldsymbol{r}') + \phi_B^*(\boldsymbol{r}')]}_{=1} \rangle_{\mathrm{EP}} \\ &\quad - \frac{R_{\mathrm{e}}^d}{\sqrt{\mathcal{N}}} \left\langle \frac{\delta \phi_A^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')} + \frac{\delta \phi_B^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')} + \frac{\delta \phi_A^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')} + \frac{\delta \phi_B^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')} \right\rangle_{\mathrm{EP}} \end{aligned} \tag{46}$$

This demonstrates that the saddle point approximation in U enforces the incompressibility constraint only on average, but the literal fluctuations of the total density in the EP theory do not vanish.

Unfortunately, the two methods for calculating composition fluctuations yield different results. Of course, after the saddle point approximation for the field U it is not a priori obvious to which extent the EP theory can correctly describe composition fluctuations. In the following we shall utilize Eqs. 25 and 26 to calculate the thermal average of the composition and its fluctuations in the EP theory. As we have argued above, these expressions will converge to the exact result if the fluctuations in U become Gaussian.

3.2 Self-Consistent Field (SCF) Theory: Saddle Point Integration in U and W

In the self-consistent field (SCF) theory one additionally approximates the functional integral over the field W by its saddle point value. Using the functionals

$$\phi_\alpha^*[W_A, W_B] = \phi_\alpha^* \left[\frac{1}{2}(\mathrm{i}U^*[W] + W), \frac{1}{2}(\mathrm{i}U^*[W] - W) \right]$$

(cf. Eq. 28), the derivative of the free energy functional $\mathcal{H}[W]$ (Eq. 42) can be written as

$$\frac{\delta\mathcal{H}[W]}{\delta W(\boldsymbol{r})} = \frac{\sqrt{\bar{\mathcal{N}}}k_{\mathrm{B}}T}{R_{\mathrm{e}}^{d}}\frac{W+\chi N(\phi_A^*-\phi_B^*)}{2\chi N} \tag{47}$$

where we have used the convention (Eq. 16). Therefore, the saddle point condition simply reads:

$$W^*(\boldsymbol{r}) = -\chi N(\phi_A^*-\phi_B^*) \tag{48}$$

where ϕ_A^* and ϕ_B^* are functionals of the fields U and W.

The two Eqs. 40 and 48 define a self-consistent set of equations for the two fields W^* and U^*. Knowing the fields W^* and U^*, one can calculate the thermal average of the microscopic composition via:

$$\langle\hat{\phi}_A-\hat{\phi}_B\rangle_{\mathrm{SCF}} = -\frac{W^*}{\chi N} = \phi_A^*-\phi_B^* \tag{49}$$

3.3 Alternative Derivation of the SCF Theory

The SCF equations are often derived in a slightly different way by additionally introducing auxiliary fields also for the microscopic densities. For completeness and further reference, we shall present this derivation here.

The starting point is the partition function, Eq. 6. Then, one proceeds to introduce collective densities Φ_A, Φ_B, fields Ω_A, Ω_B, and a pressure field U to rewrite the partition function in the form:

$$\begin{aligned}
Z &\sim \frac{1}{\prod_J n_J!}\int\prod_i \mathcal{D}[\boldsymbol{r}_i]\mathcal{P}[\boldsymbol{r}_i]\mathcal{D}\Phi_A\,\mathcal{D}\Phi_B \\
&\quad\times\exp\left[-\varrho\int \mathrm{d}^d\boldsymbol{r}\left\{\chi N\Phi_A\Phi_B\right]\delta(\Phi_A+\Phi_B-1)\delta(\Phi_A-\hat{\phi}_A)\delta(\Phi_B-\hat{\phi}_B)\right. \\
&\sim \frac{1}{\prod_J n_J!}\int\prod_i \mathcal{D}[\boldsymbol{r}_i]\mathcal{P}[\boldsymbol{r}_i]\mathcal{D}\Phi_A\mathcal{D}\Phi_B\mathcal{D}\Omega_A\mathcal{D}\Omega_B\mathcal{D}U \\
&\quad\times\exp\left[-\varrho\int \mathrm{d}^d\boldsymbol{r}\left\{\chi N\Phi_A\Phi_B\frac{\mathrm{i}U}{2}\left(\Phi_A+\Phi_B-1\right)\right\}\right] \\
&\quad\times\exp\left[-\varrho\int \mathrm{d}^d\boldsymbol{r}\left\{+\mathrm{i}\Omega_A\left(\hat{\phi}_A-\Phi_A\right)+\mathrm{i}\Omega_B\left(\hat{\phi}_B-\Phi_B\right)\right\}\right] \\
&\sim \int\mathcal{D}\Phi_A\mathcal{D}\Phi_B\mathcal{D}\Omega_A\mathcal{D}\Omega_B\mathcal{D}U\exp\left[-\frac{\mathcal{F}[\Phi_A,\Phi_B,\Omega_A,\Omega_B,U]}{k_{\mathrm{B}}T}\right]
\end{aligned} \tag{50}$$

where we have used the Fourier representation of the δ-functions. The free energy functional $\mathcal{F}[\Phi_A, \Phi_B, \Omega_A, \Omega_B, U]$ is given by:

$$\frac{\mathcal{F}[\Phi_A, \Phi_B, \Omega_A, \Omega_B, U]}{\sqrt{\mathcal{N}} k_B T (V/R_e^d)} = -\sum_J \bar{\phi}_J \ln \frac{V Q_J[\mathrm{i}\Omega_A, \mathrm{i}\Omega_B]}{n_J} \tag{51}$$
$$+ \frac{1}{V}\int \mathrm{d}^d \boldsymbol{r} \left(\chi N \Phi_A \Phi_B + \frac{\mathrm{i}U}{2}\left(\Phi_A + \Phi_B - 1\right) - \mathrm{i}\Omega_A \Phi_A - \mathrm{i}\Omega_B \Phi_B \right)$$

The saddle point approximation of the partition functions yield the SCF theory. Let us first make the saddle point approximation in the fields. Like before, we denote the values of the fields at the saddle point by an asterisk.

$$\frac{\delta \mathcal{F}}{\delta \Omega_A} = 0 \quad \Rightarrow \quad \Phi_A = -\sum_J \bar{\phi}_J \frac{1}{Q_J} V \frac{\delta Q_J}{\delta i \Omega_A} = \phi_A^* \tag{52}$$
$$\frac{\delta \mathcal{F}}{\delta \Omega_B} = 0 \quad \Rightarrow \quad \Phi_B = -\sum_J \bar{\phi}_J \frac{1}{Q_J} V \frac{\delta Q_J}{\delta i \Omega_B} = \phi_B^*$$
$$\frac{\delta \mathcal{F}}{\delta U} = 0 \quad \Rightarrow \quad \Phi_A + \Phi_B = 1$$

Note that the saddle point values, $\mathrm{i}\Omega_A^*$ and $\mathrm{i}\Omega_B^*$ are real. These equations correspond to Eq. 40. Inserting these saddle point values into the free energy functional, we obtain a functional that depends only on the composition Φ_A:

$$\frac{\mathcal{F}^*[\Phi_A]}{\sqrt{\mathcal{N}} k_B T (V/R_e^d)} \tag{53}$$
$$\equiv \frac{\mathcal{F}[\Phi_A, 1-\Phi_A, \Omega_A^*, \Omega_B^*, U^*]}{\sqrt{\mathcal{N}} k_B T (V/R_e^d)}$$
$$= -\sum_J \bar{\phi}_J \ln \frac{V Q_J[\mathrm{i}\Omega_A^*[\Phi_A], \mathrm{i}\Omega_B^*[\Phi_A]]}{n_J}$$
$$+ \frac{1}{V}\int \mathrm{d}^d \boldsymbol{r} \left\{ \chi N \Phi_A \left(1-\Phi_A\right) - \mathrm{i}\Omega_A^*[\Phi_A]\Phi_A - \mathrm{i}\Omega_B^*[\Phi_A](1-\Phi_A) \right\}$$

This functional can be used to investigate the dynamics of collective composition fluctuations (cf. Sect. 5). If we proceed to make a saddle point approximation for the collective density Φ_A, we arrive at

$$\frac{\mathcal{F}^*[\Phi_A]}{\delta \Phi_A} = 0 \quad \Rightarrow \quad \mathrm{i}(\Omega_A^* - \Omega_B^*) + \chi N(\Phi_A^* - \Phi_B^*) = 0 \tag{54}$$

which is equivalent to Eq. 48 when we identify $W^* = \mathrm{i}\Omega_A^* - \mathrm{i}\Omega_B^*$.

3.4 Numerical Techniques

Both in the SCF theory and in the EP theory the central numerical task is to solve a closed set of self-consistent equations one or many times. This is usually achieved with the following general scheme:

0. Make an initial guess for the fields U or (W, U).
1. From the fields, calculate the propagators q and $q^{\dagger}$ (Eqs. 19 and 20).
2. From the propagators, calculate a new set of fields.
3. Mix the old fields and the new fields according to some appropriate prescription.
4. Go back to (1). Continue this until the fields have converged to some value within a pre-defined accuracy.

The numerical challenge is thus twofold. First, one needs an efficient iteration prescription. Second, one needs a fast method to solve the diffusion equation.

We will begin with presenting some possible iteration prescriptions. Our list is far from complete. We will not derive the methods, nor discuss the numerical background and the numerical stability. For more information, the reader is referred to the references.

We consider a general non-linear set of equations of the form

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} \,. \tag{55}$$

The value of $\boldsymbol{x}$ after the n^{th} iteration step is denoted $\boldsymbol{x}_n$.

One of the most established methods to solve non-linear systems of equation, which has often been used to solve SCF equations, is the Newton–Raphson method [32]. In this scheme, one calculates not only $\boldsymbol{g}(\boldsymbol{x}_n)$ for a given $\boldsymbol{x}_n$, but also the whole Jacobian matrix $\boldsymbol{G}_n = \frac{\delta \boldsymbol{g}}{\delta \boldsymbol{x}'}$. The $(n+1)^{\text{th}}$ iteration is then given by

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + [1 - \boldsymbol{G}_n]^{-1}(\boldsymbol{g}(\boldsymbol{x}_n) - \boldsymbol{x}_n) \,. \tag{56}$$

If this method converges at all, i.e., if the initial guess is not too far from the solution, it usually requires very few iteration steps. Unfortunately, it has the drawback that the Jacobian must be evaluated in every iteration step. If one has m independent parameters (fields), one must solve $2m$ diffusion equations. Therefore, the method becomes inefficient if the number of degrees of freedom m, i.e., the dimension of $\boldsymbol{x}$, is very large.

To speed up the calculations, one can use the information on the Jacobian from previous steps and just iterate it. This is the idea of Broyden's method [32] and other gradient-free methods [33]. Given the Jacobian $\boldsymbol{G}_n$ for one step, the Jacobian for the next step $\boldsymbol{G}_{n+1}$ is approximated by

$$\boldsymbol{G}_{n+1} - \boldsymbol{G}_n \approx \frac{[\delta \boldsymbol{g}_n - \boldsymbol{G}_n \delta \boldsymbol{x}_n]}{(\delta \boldsymbol{x}_n)^2} \otimes \delta \boldsymbol{x}_n = \frac{[\boldsymbol{g}(\boldsymbol{x}_{n+1}) - \boldsymbol{x}_{n+1}]}{(\delta \boldsymbol{x}_n)^2} \otimes \delta \boldsymbol{x}_n \tag{57}$$

Here we have used the short notation $\delta \boldsymbol{x}_n = \boldsymbol{x}_{n+1} - \boldsymbol{x}_n$ and $\delta \boldsymbol{g}_n = \boldsymbol{g}(\boldsymbol{x}_{n+1}) - \boldsymbol{g}(\boldsymbol{x}_n)$, and $\otimes$ denotes the tensor product. The full Jacobian has to be calculated in the first step, and must possibly be updated after a number of steps. The method is a major improvement over Newton–Raphson but it still requires the evaluation of at least one Jacobian.

A very simple scheme which uses little computer time for one iteration is simple mixing. The $(n+1)^{\text{th}}$ guess $\boldsymbol{x}_{n+1}$ is simply given by

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda(\boldsymbol{g}(\boldsymbol{x}_n) - \boldsymbol{x}_n) \tag{58}$$

with some appropriate λ. Single iterations are thus cheap, but the method does not converge very well and many iteration steps are needed. The optimal values of λ have to be established empirically (typically of the order of 0.1 or less).

The simple mixing method can be improved by using information from previous steps to adjust λ. For example, the results from the two previous steps can be used to determine a new value for λ_n at every step [34]

$$\lambda_n = \sqrt{\frac{(\boldsymbol{x}_n - \boldsymbol{x}_{n-1})^2}{(\boldsymbol{g}(\boldsymbol{x}_n) - \boldsymbol{x}_n - \boldsymbol{g}(\boldsymbol{x}_{n-1}) + \boldsymbol{x}_{n-1})^2}} \,. \tag{59}$$

A more systematic approach is "Anderson mixing", which is another standard method to solve systems of non-linear equations in high dimensions [35–37]. From the remaining deviations $\boldsymbol{d}_n = \boldsymbol{g}(\boldsymbol{x}_n) - \boldsymbol{x}_n$ after n steps, one first determines

$$U_{ij} = (\boldsymbol{d}_n - \boldsymbol{d}_{n-i})(\boldsymbol{d}_n - \boldsymbol{d}_{n-j}) \tag{60}$$

$$V_j = (\boldsymbol{d}_n - \boldsymbol{d}_{n-j})\boldsymbol{d}_n \tag{61}$$

and then calculates

$$\boldsymbol{x}_n^A = \boldsymbol{x}_n + \sum_{i,j} U_{ij}^{-1} V_j (\boldsymbol{x}_{n-i} - \boldsymbol{x}_n) \tag{62}$$

$$\boldsymbol{g}_n^A = \boldsymbol{g}(\boldsymbol{x}_n) + \sum_{i,j} U_{ij}^{-1} V_j (\boldsymbol{g}(\boldsymbol{x}_{n-i}) - \boldsymbol{g}(\boldsymbol{x}_n)) \,. \tag{63}$$

In the original Anderson mixing scheme, the $(n+1)^{\text{th}}$ guess $\boldsymbol{x}_{n+1}$ is given by $\boldsymbol{x}_{n+1} = \boldsymbol{g}_n^A$. More generally, the method can be combined with simple mixing

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n^A + \lambda \left(\boldsymbol{g}_n^A - \boldsymbol{x}_n^A\right) \,. \tag{64}$$

This summarizes some methods which can be used to solve the self-consistent equations.

Regardless of the particular choice, the single most time-consuming part in every SCF or EP calculation is the repeated solution of the diffusion equation (Eqs. 19 and 20). Therefore, the next task is to find an efficient method to

solve equations of the type

$$\frac{\partial q(\boldsymbol{r},t)}{\partial t} = \frac{R_e^2}{6}\Delta q(\boldsymbol{r},t) - W(\boldsymbol{r})q(\boldsymbol{r},t)\,. \tag{65}$$

with the initial condition $q(\boldsymbol{r},0) \equiv 1$. Unfortunately, the two terms on the right hand side of this equation are incompatible. The Laplace operator Δ is best treated in Fourier space. The contribution of the field $W(\boldsymbol{r})$ can be handled much more easily in real space.

In SCF calculations of equilibrium mean-field structures, it is often convenient to operate in Fourier space, because crystal symmetries can be exploited. Relatively few Fourier components are often sufficient to characterize the structure of a phase satisfactorily. In Fourier space, Eq. 65 turns into a matrix equation

$$\frac{\partial q(\boldsymbol{k},t)}{\partial t} = -\sum_{\boldsymbol{k},\boldsymbol{k}'} \boldsymbol{A}_{\boldsymbol{k},\boldsymbol{k}'}(t) q(\boldsymbol{k}',t) \tag{66}$$

with

$$\boldsymbol{A}_{\boldsymbol{k},\boldsymbol{k}'} = \frac{R_e^2}{6}k^2\delta_{\boldsymbol{k},\boldsymbol{k}'} + W(\boldsymbol{k},\boldsymbol{k}') \tag{67}$$

and the initial condition $\boldsymbol{q}(\boldsymbol{k},0) = \delta(\boldsymbol{k},0)$. The formal solution of this equation is

$$q(\boldsymbol{k},t) = \sum_{\boldsymbol{k}'} [\exp(-\boldsymbol{A}t)]_{\boldsymbol{k},\boldsymbol{k}'} q(\boldsymbol{k}',0)\,. \tag{68}$$

It can be evaluated by diagonalizing the matrix $\boldsymbol{A}$. The approach has the advantage that the diffusion equation is solved exactly in time, i.e., one has no errors due to a time discretization. Unfortunately, the efficiency drops rapidly if the number of Fourier components and hence the dimension of the matrix $\boldsymbol{A}$ is very large. This hampers studies of fluctuations and dynamic phenomena, because one can no longer exploit crystal symmetries of the bulk structure to reduce the number of variables. It is also an issue in SCF calculations if the symmetry of the resulting structure is not known a priori [38].

Therefore, real space methods have received increased interest in recent years [38]. In real space, the number of degrees of freedom m can be increased more easily (the computing time typically scales like m or $m \ln m$). However, the time integration of the diffusion equation can no longer be performed exactly.

Several different integration methods are available in the literature. Many of them approximate the effect of the Laplace operator by finite differences [32, 39]. Among the more elaborate schemes of this kind we mention the Crank–Nicolson scheme [32] and the Dufort–Frankel scheme [39].

In our own applications, we found that a different approach was far superior to finite difference methods: The pseudo-spectral split-operator

scheme [40, 41]. This method combines Fourier space and real space steps, treating the contribution of the fields $W(\boldsymbol{r})$ in real space, and the contribution of the Laplace operator in Fourier space. In that approach, the dynamic evolution during a single discrete time step of length h is formally described by

$$q(\boldsymbol{r}, t+h) = \mathrm{e}^{-W(\boldsymbol{r})h/2}\, \mathrm{e}^{h\Delta}\, \mathrm{e}^{-W(\boldsymbol{r})h/2} q(\boldsymbol{r}, t)\,, \tag{69}$$

where the effect of the exponential $\exp(h\Delta)$ is evaluated in Fourier space. Taking as the starting point a given $q(\boldsymbol{r}, t)$, the different steps of the algorithm are

1. Calculate $q'(\boldsymbol{r}, t) = \exp(-\,W(\boldsymbol{r})h/2) q(\boldsymbol{r}, t)$ in real space for each $\boldsymbol{r}$.
2. Fourier transform $q'(\boldsymbol{r}, t) \to q'(\boldsymbol{k}, t)$.
3. Calculate $q''(\boldsymbol{k}, t) = \exp(-\,h\Delta) q'(\boldsymbol{k}, t) = \exp(-\,h(2\pi \boldsymbol{k}/L)^d) q'(\boldsymbol{k}, t)$, where d is the dimension of space and L is the system size.
4. Inverse Fourier transform $q''(\boldsymbol{k}, t) \to q''(\boldsymbol{r}, t)$.
5. Calculate $q(\boldsymbol{r}, t+h) = \exp(-\,W(\boldsymbol{r})h/2)\, q''(\boldsymbol{r}, t)$ in real space for each $\boldsymbol{r}$.
6. Go back to 1.

The algorithm thus constantly switches back and forth between the real space and the Fourier space. Even if this is done with efficient Fast Fourier Transform (FFT) routines, a single time step takes longer in this scheme than a single time step in one of the more efficient finite difference schemes. Moreover, the computing time now scales $m \log m$ with the system size, instead of m. Nevertheless, we found that this is more than compensated for by the vastly higher accuracy of the pseudo-spectral method, compared to finite difference methods. In order to calculate $q(\boldsymbol{r}, t)$ with a given numerical accuracy, the time step h could be chosen an order of magnitude larger in the pseudo-spectral method than, e.g., in the Dufort–Frankel scheme [42].

3.5 Extensions

In order to make a connection to the experiments a variety of extensions to the Gaussian chain model have been explored. Often it is desirable to include some of the architectural details for a better description on the length scale of the statistical segment length. Of course, the Gaussian chain model is only applicable if the conformational statistics on the length scale on which the composition varies is Gaussian. At strong segregation, $\chi N \gg 1$, or at surfaces to substrates or to vapor/vacuum, this assumption is often not fulfilled.

The crossover from the Gaussian chain behavior on large length scales to the rod-like behavior on the length scale of the statistical segment length can be described by the worm-like chain model [18]. In this case, the end-segment distribution q depends both on the spatial coordinate $\boldsymbol{r}$ as well as on

the orientation $\boldsymbol{u}$, defined by the tangent vector of the space curve at position $\boldsymbol{r}$ [43, 44].

While the worm-like chain model still results in an analytical diffusion equation for the propagator, the structure of a polymer in experiments or computer simulation models is often not well-described by rod-like behavior. In order to incorporate arbitrary chain architecture and incorporate details of the molecular structure on *all* length scales, the sum ($\int \mathcal{D}_1[\boldsymbol{R}]\mathcal{P}_1[\boldsymbol{R}]$) over the single chain conformations can be approximated by a partial enumeration over a large number of explicit chain conformations [19, 21, 22, 45, 46]. Ideally they have been extracted from a computer simulation of the liquid phase, or they are obtained from a molecular modelling approach. Typically 10^6–10^7 single chain conformations are employed in one-dimensional calculations. The longer the chains and the stronger the changes of the chain conformations due to the spatially inhomogeneous field, the more conformations have to be considered. The sample size should be large enough so that sufficient chains contribute significantly to the Boltzmann weight. The enumeration over the chain conformations is conveniently performed in parallel. To this end a small fraction of single chain conformations is assigned to each processor. Then, each processor calculates the Boltzmann weight of its conformations, the corresponding density profile, and the weight of its fraction of chain conformations. Subsequently, the total density profile is constructed by summing the weighted results of all processors. Typically, 64 or 128 processors are employed in parallel and a SCF calculation of a profile takes a few minutes on a CRAY T3E. As it is apparent, the detailed description of chain architecture comes at the expense of a large increase in computational demand.

Also, in addition to the zero-ranged repulsion between unlike species, other interactions can be included into the theory. In order to capture some fluid-like correlations and describe the details of packing in the vicinity of surfaces weighted density functionals have been used successfully [21].

Electrostatic interactions have been included into the theory. This is important for tuning the orientation of self-assembled structures in block copolymers [48], and also to describe biological systems [48].

4
Fluctuations

4.1
Examples of Fluctuation Effects and the Role of $\bar{N}$

From the general expression of the free energy functional (Eq. 11), we infer that the parameter combination $k_\mathrm{B}T\sqrt{\bar{\mathcal{N}}}$ sets the scale of free energy fluctuations in a volume comparable to the size of the chain extension. Generally, if $\bar{\mathcal{N}}$ is large, thermal fluctuations will not be important, and the self-consistent

field theory will provide an adequate description. Qualitatively, the quantity $\bar{\mathcal{N}}$ describes how strongly the molecules interdigitate: Fluctuation effects decrease with increasing number of interaction partners per molecule. There are, however, important exceptions where fluctuations change the qualitative behavior. Some examples of fluctuation effects in dense multicomponent mixtures shall be briefly mentioned:

(1) Unmixing transition in binary homopolymer blends.
In the vicinity of the critical point of a binary mixture one observes universal behavior, which mirrors the divergence of the correlation length of composition fluctuations. The universal behavior does not depend on the details of the system but only on the dimensionality of space and the type of order parameter. Therefore, binary polymer blends fall into the same universality class as mixtures of small molecules, metallic alloys, or the three-dimensional Ising model. In the vicinity of the critical point, $\chi_c N = 2$ for a symmetric blend [14], the difference of the composition of the two coexisting phases—the order parameter m—vanishes like $m \sim (\chi N - \chi_c N)^\beta$, where the critical exponent $\beta = 0.324$ is characteristic of the 3D Ising universality class. This behavior is, however, only observable in the close vicinity of the critical point [49]

$$t \equiv \frac{|\chi N - \chi_c N|}{\chi_c N} \ll \mathrm{Gi} \sim \frac{1}{\bar{\mathcal{N}}} \tag{70}$$

Fluctuations also shift the location of the critical point away from the prediction of the mean-field theory, $\frac{|\chi_c N - \chi_c^{\mathrm{MF}} N|}{\chi_c N} \sim \sqrt{\mathrm{Gi}} \sim \frac{1}{\sqrt{\bar{\mathcal{N}}}}$ [45, 50]. Outside the region marked by the Ginzburg criterion (Eq. 70), fluctuations do not change the qualitative behavior. In the limit of large $\bar{\mathcal{N}}$, the region where fluctuations dominate the behavior becomes small. The cross-over from Ising to mean-field behavior as a function of chain length has attracted much experimental [51] and simulational interest [52, 53].

(2) Fluctuations at the onset of microphase separation.
In case of diblock copolymers, the ordering is associated with composition waves of a finite periodicity in space. In contrast to homopolymer mixtures, not long-wavelength fluctuations are important but those with wavevectors that correspond to the periodicity of the ordering. For symmetric block copolymers one finds a fluctuation-induced first order transition (Brazovskii mechanism [54]) and the shift of the transition temperature is [55]

$$\chi_t N = \chi_t^{\mathrm{MF}} + 41.022 \bar{\mathcal{N}}^{-1/3} \quad \text{with} \quad \chi_t^{\mathrm{MF}} = 10.495 \tag{71}$$

(3) Capillary waves at the interface between two homopolymers.
If the systems contains interfaces, the local position of the interface can fluctuate. These capillary waves increase the area of the interface, and the free energy costs are proportional to the interface tension γ [56]. Long-wavelength

fluctuations are easily excited and influence measurements of the interfacial width w in experiments or computer simulations [57]. The resulting "apparent width" w

$$\frac{w^2}{R_e^2} = \frac{w_0^2}{R_e^2} + \frac{k_B T}{4\pi\gamma R_e^2} \ln \frac{L}{B_0}, \tag{72}$$

depends on the lateral length scale L on which the interface is observed. B_0 denotes a short-length scale cut-off on the order of the intrinsic width w_0 predicted by the self-consistent field theory [58]. Since γR_e^2 is proportional to $\sqrt{\bar{\mathcal{N}}}$, the capillary wave contribution to the apparent interfacial width, measured in units of R_e, scales like $1/\sqrt{\bar{\mathcal{N}}}$ [59].

(4) Formation of microemulsion-like structure in the vicinity of Lifshitz points

Capillary waves do not only broaden the width of the interface but they can also destroy the orientational order in highly swollen lamellar phases (see Fig. 1 for a phase diagram extracted from Monte Carlo Simulations). Those phases occur in mixtures of diblock-copolymers and homopolymers. The addition of homopolymers swells the distance between the lamellae, and the self-consistent field theory predicts that this distance diverges at Lifshitz points. However, general considerations show that mean-field approximations are bound to break down in the vicinity of Lifshitz points [61]. (The upper critical dimension is $d_u = 8$). This can be quantified by a Ginzburg criterion. Fluctuations are important if

$$t \equiv \frac{\chi N - \chi_t N}{\chi_t N} \ll \bar{\mathcal{N}}^{-2/5} \tag{73}$$

The Ginzburg criterion has a simple heuristic interpretation: Fluctuations of the interface position destroy the long-ranged lamellar order. Those interface fluctuations are not limited by the increase of the surface area ($\gamma = 0$), but rather by the bending stiffness κ. The correlation length ξ_n of the direction of a single, stiff interface is given by $\xi_n \sim \exp(2\pi\kappa/k_B T)$ [62]. If the stiffness κ is of order unity, there is no long-range order of the normal direction of the interface, and the lamellar order will be replaced by a microemulsion-like structure. In the vicinity of a Lifshitz point self-consistent field theory predicts the interface stiffness to vanish like $\kappa \sim k_B T \sqrt{\bar{\mathcal{N}}} t^{5/4}$ which is compatible with the Ginzburg criterion (Eq. 73) [63].

Similar arguments hold for a tricritical Lifshitz point, where a Lifshitz point happens to merge with a regular tricritical point. Here, fluctuations are important if

$$t \equiv \frac{\chi N - \chi_t N}{\chi_t N} \ll \bar{\mathcal{N}}^{-2/3} \tag{74}$$

Again this behavior is compatible with the behavior of the bending stiffness, $\kappa \sim k_B T \sqrt{\bar{\mathcal{N}}} t^{3/4}$, as obtained from self-consistent field calculations [63].

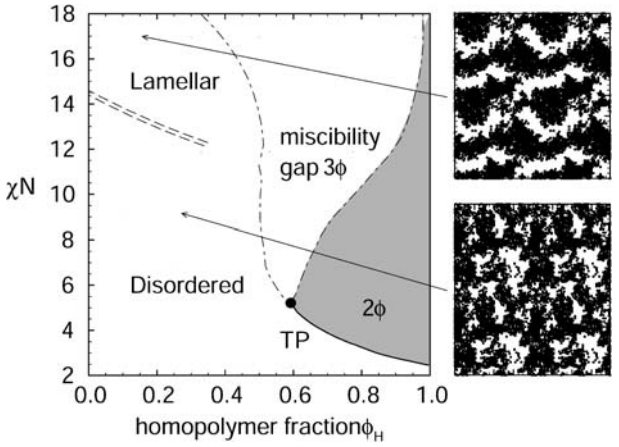

Fig. 1 Phase diagram for a ternary symmetric $A + B + AB$ blend from Monte Carlo simulations of the bond fluctuation model as a function of the incompatibility parameter χN and the homopolymer volume fraction Φ_H. The chain length of all three species is equal, $N = 32$ corresponding to $\bar{\mathcal{N}} = 91$. 2ϕ denotes a region of two-phase coexistence between an A-rich and a B-rich phase, 3ϕ one of three-phase coexistence between an A-rich, a B-rich, and a lamellar phase. Slice through the three-dimensional system of linear dimension $L = 4.7R_e$ and three periodic images are shown on the left hand side. From [60]

(5) Phase diagram of random copolymers.
Macrophase separation can also occur in random copolymers [64] which consist of a random sequence of Q blocks each of which comprises either M segments of type A or M segments of type B. Macrophase separation occurs when χM is of order unity, i.e., independent from the number of blocks, and the A-rich and B-rich phases differ in their composition only by an amount of order $1/\sqrt{Q}$. The strength of fluctuation effects can be quantified by the Ginzburg number [65]

$$\mathrm{Gi} \sim \frac{Q^2}{\bar{\mathcal{M}}} \quad \text{with} \quad \bar{\mathcal{M}} = \left(\rho_m R_M^3/M\right)^2 \tag{75}$$

where $\rho_m = N\varrho$ denotes the monomer number density. While Gi decreases in the case of homopolymer mixtures like $1/\bar{\mathcal{N}}$, it increases quadratically in the number of blocks for random copolymers. In other words, if we increase the number of blocks per molecule, fluctuation effects become stronger and stronger. Indeed, Monte Carlo simulations indicate the presence of a disordered, microemulsion-like structure in much of the parameter region where mean-field theory predicts macrophase separation.

In all but the last example, long-ranged fluctuation effects in three dimensions are controlled by the chain length N, in the sense that if $N \sim \bar{\mathcal{N}} \to \infty$, fluctuation effects can be made arbitrarily small. This provides a formal limit in which self-consistent field theory becomes accurate. We emphasize, how-

ever, that in many practical examples fluctuations are very important. For instance, the broadening of the interface width by capillary waves is typically a sizeable fraction of the intrinsic width calculated by self-consistent field theory. [1]

4.2 Gaussian Fluctuations and Random Phase Approximation

As long as fluctuations are weak, they can be treated within the Gaussian approximation. One famous example of such a treatment is the random phase approximation (RPA) which describes Gaussian fluctuations in homogeneous phases. The RPA has been extended to inhomogeneous saddle points by Shi, Noolandi, and coworkers [66, 67]. In this section, we shall re-derive this generalized RPA theory within the formalism developed in the previous sections.

Our starting point is the Hamiltonian $\mathcal{G}[U, W]$ of Eq. 11. We begin with expanding it up to second order about the saddle point $\bar{\mathcal{G}} = \mathcal{G}[U^*, W^*]$ of the SCF theory. To this end, we define the averaged single chain correlation functions [13, 66, 67]

$$K_{\alpha\beta}(\boldsymbol{r}, \boldsymbol{r}') = -\frac{\delta\phi^*_\alpha}{\delta W_\beta} = \frac{1}{\varrho}\sum_J n_J \frac{\delta^2 \ln VQ_J[W_A, W_B]}{\delta W_\alpha(\boldsymbol{r})\delta W_\beta(\boldsymbol{r}')} \tag{76}$$

with $\alpha, \beta = A$ or B, where ϕ^*_α is defined as in Eq. 28, and all derivatives are evaluated at the saddle point values of the fields, $W^*_A = (\mathrm{i}U^* + W^*)/2$ and $W^*_B = (\mathrm{i}U^* - W^*)/2$). As usual the sum J runs over the different types of polymers in the mixture, and n_J denotes the total number of polymers of type J. In the case of a homopolymer mixture, the mixed correlation functions $K_{AB}(\boldsymbol{r}, \boldsymbol{r}')$, $K_{BA}(\boldsymbol{r}, \boldsymbol{r}')$ are zero.

With these definitions, the quadratic expansion of $\mathcal{G}[U, W]$ (Eq. 11) can be written as

$$\begin{aligned}\frac{\mathcal{G}[U, W] - \bar{\mathcal{G}}}{\sqrt{\mathcal{N}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} = &\frac{1}{V}\frac{1}{4\chi N}\int \mathrm{d}^d\boldsymbol{r}\,\delta W(\boldsymbol{r})^2 \\ &- \frac{1}{2V}\sum_{\alpha\beta}\int \mathrm{d}^d\boldsymbol{r}\,\mathrm{d}^3\boldsymbol{r}' K_{\alpha\beta}(\boldsymbol{r}, \boldsymbol{r}')\delta W_\alpha(\boldsymbol{r})\delta W_\beta(\boldsymbol{r}') \\ &\text{with} \quad \delta W_A = (\mathrm{i}\delta U + \delta W)/2\,, \quad \delta W_B = (\mathrm{i}\delta U - \delta W)/2\,.\end{aligned} \tag{77}$$

Since $\bar{\mathcal{G}}$ is an extremum, the linear terms in δU and δW vanish. Next we expand this expression in δU and δW. To simplify the expressions, we follow

[1] In particle-based models, the mean-field approximation also neglects short-ranged correlations in the polymeric fluid, e.g., the fluid-like packing of the particles or the correlations due to the self-avoidance of the polymers on short length scales. There is no small parameter which controls the magnitude of these correlation effects but they are incorporated into the coarse-grained parameters R_e and χN.

Laradji et al. and define [13, 67]

$$\Sigma = K_{AA} + K_{AB} + K_{BA} + K_{BB} \tag{78}$$
$$C = K_{AA} - K_{AB} - K_{BA} + K_{BB} \tag{79}$$
$$\Delta = K_{AA} + K_{AB} - K_{BA} - K_{BB} \tag{80}$$
$$\Delta^+ = K_{AA} - K_{AB} + K_{BA} - K_{BB}\,, \tag{81}$$

and adopt the matrix notation

$$\delta A \cdot K \cdot \delta B \equiv \int \mathrm{d}^d\boldsymbol{r}\mathrm{d}^d\boldsymbol{r}' K(\boldsymbol{r},\boldsymbol{r}')\delta A(\boldsymbol{r})\delta B(\boldsymbol{r}')\,.$$

with the unity operator $1 = \delta(\boldsymbol{r} - \boldsymbol{r}')$. Note that the operators Σ and C are symmetric, and Δ and Δ^+ are transposed to each other. In the new notation, Eq. 77 reads

$$\frac{\mathcal{G}[U,W] - \bar{\mathcal{G}}}{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} = \frac{1}{V}\frac{1}{4\chi N}\delta W 1 \delta W \tag{82}$$
$$+ \frac{1}{8V}\left(\delta U \Sigma \delta U - \delta W C \delta W + \mathrm{i}\delta W \Delta \delta U + \mathrm{i}\delta U \Delta^+ \delta W\right)\,.$$

and can, after some algebra, be cast in the form

$$\frac{\mathcal{G}[U,W] - \bar{\mathcal{G}}}{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} = \tag{83}$$
$$\frac{1}{8V}\left\{(\delta U - \delta U^*)\Sigma(\delta U - \delta U^*) + \frac{2}{\chi N}\delta W[1 - \frac{\chi N}{2}\tilde{C}]\delta W\right\}$$
$$\text{with}\quad \tilde{C} = C - \Delta\Sigma^{-1}\Delta^+ \quad\text{and}\quad \delta U^* = -\,\mathrm{i}\Sigma^{-1}\Delta^+\delta W\,. \tag{84}$$

The field $U^* + \delta U^*$ corresponds to the saddle point of $\mathcal{G}[U,W]$ at fixed $W = W^* + \delta W$.

Both the fluctuations $\langle \delta W^2\rangle$ and $\langle(\delta U - \delta U^*)^2\rangle$ decrease like $1/\sqrt{\bar{\mathcal{N}}}$. However, the fluctuations in both fields are qualitatively different. The operator Σ is positive, i.e., the eigenvalues of Σ are positive, therefore fluctuations in U remain small and the Gaussian approximation is justified in the limit of large $\sqrt{\bar{\mathcal{N}}}$. The operator $1 - \chi N/2\tilde{C}$, on the other hand, may have negative eigenvalues at large χ. In that case the saddle point under consideration is unstable and the true solution of the SCF theory corresponds to a different structure (phase). A saddle point which turns from being stable to unstable defines a spinodal or a continuous phase transition. In addition, the operator $1 - \chi N/2\tilde{C}$ also has zero eigenvalues in the presence of continuous symmetries. In all of these cases, the expansion for small δW becomes inaccurate and fluctuations in δW may give rise to qualitative deviations from the mean-field theory.

Within the Gaussian approximation, the average $\langle W(\boldsymbol{r})W(\boldsymbol{r}')\rangle$ can readily be calculated:

$$\langle W(\boldsymbol{r})W(\boldsymbol{r}')\rangle = W^*(\boldsymbol{r})W^*(\boldsymbol{r}') + \frac{2\chi N}{\sqrt{\bar{\mathcal{N}}}} R_{\mathrm{e}}^d \left[1 - \frac{\chi N}{2}\tilde{C}\right]^{-1}_{\boldsymbol{r},\boldsymbol{r}'} \tag{85}$$

Using the relation between fluctuations of W and the composition, Eq. 26, we obtain

$$\begin{aligned}
&\langle[\hat{\phi}_A(\boldsymbol{r}) - \hat{\phi}_B(\boldsymbol{r})][\hat{\phi}_A(\boldsymbol{r}') - \hat{\phi}_B(\boldsymbol{r}')]\rangle - \langle\hat{\phi}_A(\boldsymbol{r}) - \hat{\phi}_B(\boldsymbol{r})\rangle\langle\hat{\phi}_A(\boldsymbol{r}') - \hat{\phi}_B(\boldsymbol{r}')\rangle \qquad (86)\\
&\quad\equiv \frac{2R_{\mathrm{e}}^d}{\sqrt{\bar{\mathcal{N}}}\chi N}\left[\left[1 - \frac{\chi N}{2}\tilde{C}\right]^{-1} - 1\right]_{\boldsymbol{r},\boldsymbol{r}'}\\
&\quad= \frac{2R_{\mathrm{e}}^d}{\sqrt{\bar{\mathcal{N}}}\chi N}\left[\sum_{k=1}^{\infty}\left[\frac{\chi N}{2}\tilde{C}\right]^k\right]_{\boldsymbol{r},\boldsymbol{r}'} = \frac{1}{\sqrt{\bar{\mathcal{N}}}}\left[\tilde{C}\left[1 - \frac{\chi N}{2}\tilde{C}\right]\right]^{-1}_{\boldsymbol{r},\boldsymbol{r}'}\\
&\quad= \frac{1}{\sqrt{\bar{\mathcal{N}}}}\left[\tilde{C}^{-1} - \frac{\chi N}{2}1\right]^{-1}_{\boldsymbol{r},\boldsymbol{r}'}.
\end{aligned}$$

This is the general expression for Gaussian composition fluctuations in incompressible polymer blends derived from EP theory. The original derivation of Shi, Noolandi and coworkers [66, 67] uses as a starting point the density functional (Eq. 53) and gives the identical result. A generalization for compressible blends can be found in [13].

The special situation of a homogeneous saddle point, corresponding to a homogeneous disordered phase, is particularly interesting. In that case, explicit analytical relations between the single chain partition function and the fields can be obtained, and one recovers the well-known random phase approximation (RPA). To illustrate this approach, we shall now derive the RPA structure factor for the case of a symmetric binary homopolymer blend.

Since the reference saddle point is spatially homogeneous, all correlation functions $K_{\alpha\beta}(\boldsymbol{r},\boldsymbol{r}')$ only depend on $(\boldsymbol{r}-\boldsymbol{r}')$ and it is convenient to perform the calculations in Fourier space. We use the following conventions for the Fourier expansion:

$$\hat{\phi}(\boldsymbol{r}) = \bar{\phi} + \sum_{\boldsymbol{q}\neq 0}\phi_{\boldsymbol{q}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{q}\boldsymbol{r}} \qquad \phi_{\boldsymbol{q}} = \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\hat{\phi}(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\boldsymbol{r}}\,. \tag{87}$$

In homopolymer melts, the mixed single chain correlation functions K_{AB}, K_{BA} vanish, thus one has $C = \Sigma = K_{AA} + K_{BB}$ and $\Delta = \Delta^+ = K_{AA} - K_{BB}$. An explicit expression for the functional $\phi_A^*(\boldsymbol{r}) = \bar{\phi}_A\langle\hat{\phi}_A(\boldsymbol{r})\rangle$ in the limit of small

fields W_A can be derived:

$$\phi_A^*(\boldsymbol{r}) = \bar{\phi}_A \left(1 - \sum_{\boldsymbol{q}\neq 0} g_A(\boldsymbol{q}) W_{A\boldsymbol{q}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\boldsymbol{r}}\right) \tag{88}$$
$$= \bar{\phi}_A \left(1 - \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}' \sum_{q\neq 0} g_A(\boldsymbol{q}) W_A(\boldsymbol{r}')\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}(\boldsymbol{r}-\boldsymbol{r}')}\right)$$

A similar expression is valid for $\phi_B^*(\boldsymbol{r})$. Here $g(\boldsymbol{q}) = (2/x^2)(\exp[-x] - 1 + x)$ with $x \equiv R_\mathrm{e}^2\boldsymbol{q}^2/6$ denotes the Debye function. Hence, the Fourier transform of K_{AA} can be identified with the single chain structure factor:

$$K_{AA}(\boldsymbol{r},\boldsymbol{r}') = -\frac{\delta\phi_A^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')} = \bar{\phi}_A \frac{1}{V}\sum_{\boldsymbol{q}\neq 0} g_A(\boldsymbol{q})\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}(\boldsymbol{r}-\boldsymbol{r}')} \tag{89}$$
$$K_{AA}(\boldsymbol{q}) = \bar{\phi}_A\, g_A(\boldsymbol{q})\,,$$

and a similar expression holds K_{BB}. Using the definition of $\tilde{C}$, Eq. 84, one obtains

$$\tilde{C}^{-1} = \frac{1}{4}\left[\frac{1}{K_{AA}} + \frac{1}{K_{BB}}\right] = \frac{1}{4}\left[\frac{1}{\bar{\phi}_A\, g_A} + \frac{1}{\bar{\phi}_B g_B}\right].$$

Inserting this expression into Eq. 86, we can also obtain an explicit expression for the EP Hamiltonian $\mathcal{H}[W] \equiv \mathcal{G}[U^*, W]$ of binary blends within the RPA:

$$\frac{\mathcal{H}_\mathrm{RPA}[W] - \mathcal{H}^{\bar{}}}{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T(V/R_\mathrm{e}^d)} = +\frac{1}{8}\sum_{\boldsymbol{q}\neq 0}\left(\frac{2}{\chi N} - \frac{4\bar{\phi}_A g_A(\boldsymbol{q})\bar{\phi}_B g_B(\boldsymbol{q})}{\bar{\phi}_A g_A(\boldsymbol{q}) + \bar{\phi}_B g_B(\boldsymbol{q})}\right)|W_{\boldsymbol{q}}|^2 \tag{90}$$

Combined with Eq. 26, this yields the well-known RPA expression for the collective structure factor in binary polymer blends

$$\frac{4}{\sqrt{\bar{\mathcal{N}}}(V/R_\mathrm{e}^d)\langle|\phi_{A\boldsymbol{q}} - \phi_{B\boldsymbol{q}}|^2\rangle} = \frac{1}{\bar{\phi}_A\, g_A(\boldsymbol{q})} + \frac{1}{\bar{\phi}_B g_B(\boldsymbol{q})} - 2\chi N\,. \tag{91}$$

This calculation also justifies the use of Eq. 26 to calculate the fluctuations in the EP theory [29]. If we used the literal fluctuations in the EP theory according of Eq. 44, we would not recover the RPA expression but rather [29]

$$\langle|\hat{\phi}_{A\boldsymbol{q}} - \hat{\phi}_{B\boldsymbol{q}}|^2\rangle_\mathrm{EP} \tag{92}$$
$$= \langle|\phi_{A\boldsymbol{q}}^* - \phi_{B\boldsymbol{q}}^*|^2\rangle - \frac{R_\mathrm{e}^d}{\sqrt{\bar{\mathcal{N}}}V^2}\int \mathrm{d}^d\boldsymbol{r}\mathrm{d}^3\boldsymbol{r}'\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}(\boldsymbol{r}-\boldsymbol{r}')}\left\langle\frac{\delta\phi_A^*(\boldsymbol{r})}{\delta W_A(\boldsymbol{r}')} + \frac{\delta\phi_B^*(\boldsymbol{r})}{\delta W_B(\boldsymbol{r}')}\right\rangle$$
$$= \frac{8\chi N g^2(\boldsymbol{q})\bar{\phi}_A^2\bar{\phi}_B^2}{\sqrt{\bar{\mathcal{N}}}(V/R_\mathrm{e}^d)(1 - 2\chi N\bar{\phi}_A\bar{\phi}_B g(\boldsymbol{q}))} + \frac{g(\boldsymbol{q})}{\sqrt{\bar{\mathcal{N}}}(V/R_\mathrm{e}^d)}$$
$$= \langle|\hat{\phi}_{A\boldsymbol{q}} - \hat{\phi}_{B\boldsymbol{q}}|^2\rangle + \frac{(1 - 4\bar{\phi}_A\bar{\phi}_B)g(\boldsymbol{q})}{\sqrt{\bar{\mathcal{N}}}(V/R_\mathrm{e}^d)}\,.$$

Both contributions are of the same order.

This example illustrates how the RPA can be used to derive explicit, analytical expressions for Hamiltonian structure factors. The generalization to blends that also contain copolymers is straightforward.

4.3 Relation to Ginzburg–Landau Models

In order to make the connection to the Landau–Ginzburg theory for binary blends, we study the behavior of the structure factor at small wavevectors $\boldsymbol{q}$ for a symmetric mixture. Using $g(\boldsymbol{q}) \approx 1 + \frac{(qR_e)^2}{18} + \cdots$ we obtain:

$$\frac{4}{\sqrt{\bar{\mathcal{N}}}(V/R_e^d)\langle|\phi_{A\boldsymbol{q}} - \phi_{B\boldsymbol{q}}|^2\rangle} = \frac{1}{\bar{\phi}_A} + \frac{1}{\bar{\phi}_B} - 2\chi N + \frac{\boldsymbol{q}^2 R_e^2}{18\bar{\phi}_A} + \frac{\boldsymbol{q}^2 R_e^2}{18\bar{\phi}_B} \tag{93}$$

If we assume composition fluctuations to be Gaussian, we can write down a free energy functional compatible with Eq. 91.

$$\begin{aligned}\frac{\mathcal{F}_{\mathrm{RPA}}[\phi_{A\boldsymbol{q}}]}{\sqrt{\bar{\mathcal{N}}}k_B T(V/R_e^d)} &= \frac{1}{\sqrt{\bar{\mathcal{N}}}(V/R_e^d)}\frac{(2\phi_{A\boldsymbol{q}} - 1)^2}{2\langle|\phi_{A\boldsymbol{q}} - \phi_{B\boldsymbol{q}}|^2\rangle} \\ &= \frac{1}{2}\left(\frac{1}{\bar{\phi}_A} + \frac{1}{1-\bar{\phi}_A} - 2\chi N\right)|\phi_{A0} - \bar{\phi}_A|^2 \\ &\quad + \frac{1}{2}\sum_{q\neq 0}\left(\frac{1}{\bar{\phi}_A} + \frac{1}{1-\bar{\phi}_A} - 2\chi N\right. \\ &\quad \left. + \frac{\boldsymbol{q}^2 R_e^2}{18\bar{\phi}_A} + \frac{\boldsymbol{q}^2 R_e^2}{18(1-\bar{\phi}_A)}\right)|\phi_{A\boldsymbol{q}}|^2\end{aligned} \tag{94}$$

Note that this free energy functional is Gaussian in the Fourier coefficients of the composition, and, hence, the critical behavior still is of mean-field type. Transforming back from Fourier expansion for the spatial dependence to real space, we obtain for the free energy functional:

$$\begin{aligned}\frac{\mathcal{F}_{\mathrm{RPA}}[\phi_A(\boldsymbol{r})]}{\sqrt{\bar{\mathcal{N}}}k_B T(V/R_e^d)} &= \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\left\{\frac{1}{2}\left(\frac{1}{\bar{\phi}_A} + \frac{1}{1-\bar{\phi}_A} - 2\chi N\right)\left(\phi_A(\boldsymbol{r}) - \frac{1}{2}\right)^2\right. \\ &\quad \left. + \frac{R_e^2}{36\bar{\phi}_A(1-\bar{\phi}_A)}(\nabla\phi_A)^2\right\} \\ &= \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\left\{\frac{1}{2}\frac{\mathrm{d}^2 f_{\mathrm{FH}}}{\mathrm{d}\bar{\phi}_A^2}\left(\phi_A(\boldsymbol{r}) - \frac{1}{2}\right)^2 + \frac{R_e^2(\nabla\phi_A)^2}{36\bar{\phi}_A(1-\bar{\phi}_A)}\right\}\end{aligned} \tag{95}$$

Expanding the free energy of the homogeneous system (Flory–Huggins free energy of mixing), we can restore higher order terms and obtain the Landau–

de Gennes free energy functional for a symmetric binary polymer blend:

$$\frac{\mathcal{F}_{\mathrm{GL}}[\phi_A(\boldsymbol{r})]}{\sqrt{\mathcal{N}}k_{\mathrm{B}}T(V/R_{\mathrm{e}}^d)} = \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\left\{f_{\mathrm{FH}}(\phi_A) + \frac{R_{\mathrm{e}}^2}{36\bar{\phi}_A(1-\bar{\phi}_A)}(\nabla\phi_A)^2\right\} \tag{96}$$

$$= \frac{1}{V}\int \mathrm{d}^d\boldsymbol{r}\left\{(2-\chi N)\left(\phi_A(\boldsymbol{r})-\frac{1}{2}\right)^2 + \frac{4}{3}\left(\phi_A(\boldsymbol{r})-\frac{1}{2}\right)^4 + \frac{R_{\mathrm{e}}^2}{36\bar{\phi}_A(1-\bar{\phi}_A)}(\nabla\phi_A)^2\right\}$$

A field-theory based on this simple expansion will already yield three-dimensional Ising critical behavior. Note that the non-trivial critical behavior is related to fluctuations in W or $\phi_A - \phi_B$. Fluctuations in the incompressibility field U or the total density $\phi_A + \phi_B$ are not important.

4.4 Field-Theoretic Polymer Simulations

Studying fluctuations beyond the Gaussian approximation is difficult. Special types of fluctuations, e.g., capillary waves of interfaces, can sometimes be described analytically within suitable approximations [58]. The only truly general methods are however computer simulations. Here we shall discuss two different approaches to simulating field theories for polymers: Langevin simulations and Monte Carlo simulations.

4.4.1 Langevin Simulations

As long as one is mainly interested in composition fluctuations (EP approximation, see Sect. 3), the problem can be treated by simulation of a real Langevin process. The correct ensemble is reproduced by the dynamic equation

$$\frac{\partial W(\boldsymbol{r},t)}{\partial t} = -\int \mathrm{d}^d\boldsymbol{r}' M(\boldsymbol{r},\boldsymbol{r}')\frac{\delta\mathcal{H}[W]}{\delta W(\boldsymbol{r}',t)} + \theta(\boldsymbol{r},t)\,, \tag{97}$$

where $M(\boldsymbol{r},\boldsymbol{r}')$ is an (arbitrary) kinetic coefficient, and $\theta(\boldsymbol{r},t)$ is a stochastic noise. The first two moments of θ are fixed by the fluctuation-dissipation theorem

$$\langle\theta\rangle = 0 \quad \langle\theta(\boldsymbol{r},t)\theta(\boldsymbol{r}',t)\rangle = 2k_{\mathrm{B}}TM(\boldsymbol{r},\boldsymbol{r}')\delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t-t')\,. \tag{98}$$

The choice of $M(\boldsymbol{r},\boldsymbol{r}')$ determines the dynamic properties of the system. For example, $M(\boldsymbol{r},\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}')/\tau k_{\mathrm{B}}T$ corresponds to a non-conserved field, while field conservation can be enforced by using a kinetic coefficient of the form $M(\boldsymbol{r},\boldsymbol{r}') = \nabla_{\boldsymbol{r}}\Lambda(\boldsymbol{r}-\boldsymbol{r}')\nabla_{\boldsymbol{r}'}$. Different forms for the Onsager coefficient Λ will be discussed in Sect. 5. In each step of the Langevin simulation

one updates all field variables simultaneously [29] and the the self-consistent equations for the saddle point $U^*(W)$ have to be solved.

The approach is commonly referred to as external potential dynamics (EPD). A related approach was originally introduced by Maurits and Fraaije [31]. However, these authors do not determine $U^*(W)$ exactly, but only approximately by solving separate Langevin equations for real fields W_A and W_B. This amounts to introducing a separate Langevin equation for a real field iU (i.e., an imaginary U in our notation) in addition to Eq. 97.

As the fluctuation effects discussed earlier (cf. Sect. 4.1) are essentially caused by composition fluctuations, the EP approximation seems reasonable. Nevertheless, one would also like to study the full fluctuating field theory by computer simulations. In attempting to do so, however, one faces a serious problem: Since the field iU is imaginary, the Hamiltonian $\mathcal{G}$ is in general *complex*, and the "weight" factor $\exp[-\mathcal{G}]$ can no longer be used to generate a probability function. This is an example of a sign problem, as is well-known from other areas of physics [68], e.g., lattice gauge theories and correlated fermion theories. A number of methods have been devised to handle complex Hamiltonians or complex actions [69–73]. Unfortunately, none of them is as universally powerful as the methods that sample real actions (Langevin simulations, Monte Carlo, etc.).

Fredrickson and coworkers [74–77] have recently introduced the complex Langevin method [69, 70] into the field of polymer science. The idea of this method is simply to extend the real Langevin formalism, e.g., as used in EPD, to the case of a complex action. The drift term $\delta\mathcal{H}/\delta W$ in Eq. 97 is replaced by complex drift terms $\delta\mathcal{G}[U, W]/\delta U$ and $\delta\mathcal{G}[U, W]/\delta W$, which govern the dynamics of complex fields U and W. This generates a diffusion process in the entire complex plane for both U and W. One might wonder why such a process should sample line integrals over U and W. To understand that, we recall that the integration paths of the line integrals can be distorted arbitrarily in the complex plane, as long as no pole is crossed, without changing the result. Hence a complex Langevin trajectory samples an ensemble of possible line integrals. Under certain conditions, the density distribution converges towards a stationary distribution which indeed reproduces the desired expectation values [78]. Unfortunately, these conditions are not generally satisfied, and cases have been reported where a complex Langevin process did not converge at all, or converged towards the wrong distribution [71, 79]. Thus the theoretical foundations of this method are still under debate. In the context of polymer simulations, however, no problems with the convergence or the uniqueness of the solution were reported.

In our case, the complex Langevin equations that simulate the Hamiltonian Eq. 11 read

$$\frac{\partial U(\boldsymbol{r},t)}{\partial t} = -\frac{\delta \mathcal{G}[U,W]/k_{\mathrm{B}}T}{\delta U(\boldsymbol{r},t)} + \theta_U(\boldsymbol{r},t) \tag{99}$$

$$\frac{\partial W(\boldsymbol{r},t)}{\partial t} = -\frac{\delta \mathcal{G}[U,W]/k_{\mathrm{B}}T}{\delta W(\boldsymbol{r},t)} + \theta_W(\boldsymbol{r},t)\,, \tag{100}$$

where $\theta_U(\boldsymbol{r},t)$, $\theta_W(\boldsymbol{r},t)$ are *real* Gaussian white noises which satisfy the fluctuation-dissipation theorem

$$\langle\theta\rangle = 0 \quad \langle\theta(\boldsymbol{r},t)\theta(\boldsymbol{r}',t)\rangle = 2\delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t-t')\,. \tag{101}$$

The partial derivatives are given by

$$\frac{\delta \mathcal{G}[U,W]/k_{\mathrm{B}}T}{\delta W(\boldsymbol{r},t)} = \frac{\sqrt{\bar{\mathcal{N}}}}{2R_{\mathrm{e}}^d}\left[\frac{W}{\chi N} + \phi_A^* - \phi_B^*\right] \tag{102}$$

$$\frac{\delta \mathcal{G}[U,W]/k_{\mathrm{B}}T}{\delta U(\boldsymbol{r},t)} = \frac{\sqrt{\bar{\mathcal{N}}}}{2R_{\mathrm{e}}^d}\left[\phi_A^* + \phi_B^* - 1\right]\,, \tag{103}$$

with ϕ_A^* and ϕ_B^* defined as in Eq. 28.

If the noise term is turned off, the system is driven towards the nearest saddle point. Therefore, the same set of equations can be used to find and test mean-field solutions. The complex Langevin method was first applied to dense melts of copolymers [74], and later to mixtures of homopolymers and copolymers [80] and to diluted polymers confined in a slit under good solvent conditions [77]. Figure 2 shows examples of average "density" configurations $\langle\phi^*\rangle$ for a ternary block copolymer/homopolymer system above and below the order/disorder transition.

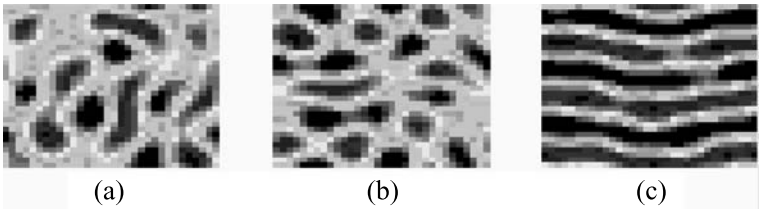

Fig. 2 Averaged densities across the order-disorder transition in a two-dimensional ternary system with *A*, *B* homopolymers and *A*–*B* copolymers (20% homopolymer volume fraction), as obtained from Complex Langevin simulation runs

4.4.2 Field-Theoretic Monte Carlo

As an alternative approach to sampling the fluctuating field theory, Düchs et al. [42, 80] have proposed a Monte Carlo method. Since the weight $\exp[-\mathcal{G}]$

is not positive semidefinite, the Monte Carlo algorithm cannot be applied directly. To avoid this problem, we split $\mathcal{G}$ into a real and an imaginary part $\mathcal{G}^R$ and $i\mathcal{G}^I$ and sample only the real contribution $\exp(-\mathcal{G}^R)$ [71]. The imaginary contribution is incorporated into a complex reweighting factor $\exp(-\mathrm{i}\mathcal{G}^I)$. Statistical averages over configurations j then have to be computed according to

$$\langle A\rangle = \frac{\sum_j \exp(-\mathrm{i}G_j^I)A_j}{\sum_j \exp(-\mathrm{i}G_j^I)}\,, \tag{104}$$

i.e., every configuration is weighted with this factor. Furthermore, we premise that the (real) saddle point $\mathrm{i}U^*$ contributes substantially to the integral over the imaginary field $\mathrm{i}U$ and shift the integration path such that it passes through the saddle point. The field $\mathrm{i}U$ is then represented as

$$\mathrm{i}U = \mathrm{i}U^*[W] + \mathrm{i}\omega\,, \tag{105}$$

where ω is real.

The Monte Carlo simulation includes two different types of moves: trial moves of W and ω. The moves in ω are straightforward. Implementing the moves in W is more involved. Every time that W is changed, the new saddle point $\mathrm{i}U^*[W]$ must be evaluated. In an incompressible blend, the set of self consistent equations

$$\phi_A^*(\boldsymbol{r}) + \phi_B^*(\boldsymbol{r}) = 1\,, \quad \forall \boldsymbol{r} \tag{106}$$

must be solved for $\mathrm{i}U^*$. Fortunately, this does not require too many iterations, if W does not vary very much from one Monte Carlo step to another.

Compared to the Complex Langevin method, the Monte Carlo method has the advantage of being well founded theoretically. However, it can become very inefficient when $\mathcal{G}^I$ spreads over a wide range and the reweighting factor oscillates strongly. In practice, it relies on the fact that the integral is indeed dominated by one (or several) saddle points.

Can we expect this to be the case here? To estimate the range of the reweighting factor, we briefly re-inspect the Hubbard–Stratonovich transformation of the total density $\hat{\phi}_A + \hat{\phi}_B$ that leads to the fluctuating field U. For simplicity we consider a one-component system. In a compressible polymer solution or blend, the contribution of the repulsive interaction energy to the partition function can be written as

$$\exp\left[-\varrho N\frac{\kappa}{2}\int \mathrm{d}^d\boldsymbol{r}(\hat{\phi}_A + \hat{\phi}_B - \phi_0)^2\right] \tag{107}$$

(with $\phi_0 = 1$ in a melt, and $\phi_0 = 0$ in a good solvent). The Hubbard–Stratonovich transformation of this expression is proportional to

$$\int \mathcal{D}U \exp\left[-\varrho\left\{\frac{1}{8\kappa N}\int \mathrm{d}^d\boldsymbol{r}U^2 + \mathrm{i}\int \mathrm{d}^d\boldsymbol{r}\frac{U}{2}(\hat{\phi}_A + \hat{\phi}_B - \phi_0)\right\}\right] \tag{108}$$

Thus U should to be distributed around the saddle point with a width proportional to κ. The method should work best for very compressible solutions with small κ. In contrast, in an incompressible blend, the contribution (Eq. 107) is replaced by a delta function constraint $\delta(\hat{\phi}_A + \hat{\phi}_B - 1)$. The fluctuating field representation of this constraint is

$$\int \mathcal{D}U \exp\left[-\mathrm{i}\varrho \int \mathrm{d}^d\boldsymbol{r}\frac{U}{2}(\hat{\phi}_A + \hat{\phi}_B - 1)\right] . \tag{109}$$

The only forces driving U towards the saddle point are now those related to the single chain fluctuations (cf. Sect. 4.2), and U will be widely distributed on the imaginary axis. Thus it is not clear whether the Monte Carlo method will work for incompressible systems.

Indeed, simulations of an incompressible ternary A+B+AB homopolymer/copolymer blend showed that the values of the argument $\mathcal{G}^I$ of the reweighting factor cover a wide range (Fig. 3). Hence computing statistical averages becomes very difficult, because both the numerator and the denominator in Eq. 104 are subject to very large relative errors.

Fortunately, a closer inspection of the data reveals that the reweighting factor is entirely uncorrelated with the fluctuating field W. Since the latter determines all quantities of interest to us (Eqs. 25 and 26), the averages can be computed without reweighting. Moreover, the time scales for variations of W and ω are decoupled. By sampling ω more often than W, they can be chosen such that the time scale of ω is much shorter than that of W. Finally, the dynamics of the actual simulation, which is governed by the real factor $\exp(-\mathcal{G}^R)$, is virtually the same as that of a reference simulation with ω switched off [80].

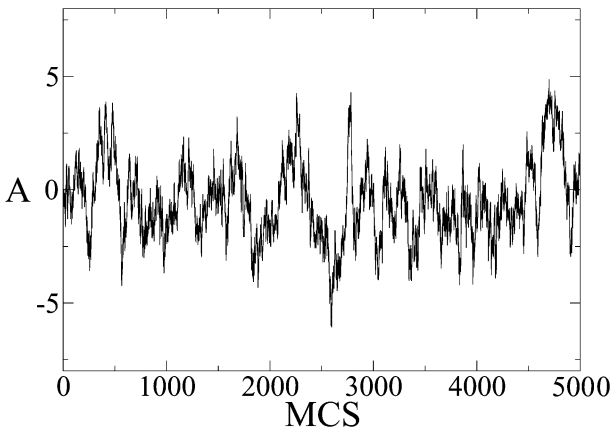

Fig. 3 Argument G^I of the reweighting factor in a simulation of an incompressible blend. From [80]

Fig. 4 Snapshots of ϕ_A^* for the same ternary $A + B + AB$ system as in Fig. 2, at 70% homopolymer volume fraction, for different χN above and below the order-disorder transition, as obtained from Monte Carlo simulations with ω switched off (EP theory). From [83]

Combining all these observations, we infer that the fluctuations in ω do not influence the composition structure of the blend substantially. A similar decoupling between composition and density fluctuations is suggested by other studies [45, 81, 82]. The composition structure can equivalently be studied in a Monte Carlo simulation which samples only W and sets $\omega \equiv 0$. Figure 4 shows examples of snapshots for the ternary system in the vicinity of the order-disorder transition.

We have demonstrated this for one specific case of an incompressible blend and suspect that it may be a feature of incompressible blends in general. The observation that the fluctuations in W and ω are not correlated with each other is presumably related to the fact that the (vanishing) density fluctuations do not influence the composition fluctuations. If that is true, we can conclude that field-theoretic Monte Carlo can be used to study fluctuations in polymer mixtures in the limits of high and low compressibilities. Whether it can also be applied at intermediate compressibilities will have to be explored in the future.

We should note that the Monte-Carlo simulation with $\omega \equiv 0$ effectively samples the EP Hamiltonian. This version of field-theoretic Monte Carlo is equivalent to the real Langevin method (EPD), and can be used as an alternative. Monte Carlo methods are more versatile than Langevin methods, because an almost unlimited number of moves can be invented and implemented. In our applications, the W and ω-moves simply consisted of random increments of the local field values, within ranges that were chosen such that the Metropolis acceptance rate was roughly 35%. In principle, much more sophisticated moves are conceivable, e.g., collective moves or combined EPD/Monte Carlo moves (hybrid moves [84]). On the other hand, EPD is clearly superior to Monte Carlo when dynamic properties are studied. This will be discussed in the next section.

5
Dynamics

5.1
Onsager Coefficients and Dynamic SCF Theory (DSCFT)

In order to describe the diffusive dynamics of composition fluctuations in binary mixtures one can extend the time-dependent Ginzburg–Landau methods to the free energy functional of the SCF theory. The approach relies on two ingredients: a free energy functional that accurately describes the chemical potential of a spatially inhomogeneous composition distribution out of equilibrium and an Onsager coefficient that relates the variation of the chemical potential to the current of the composition.

We implicitly assume that the out of equilibrium configuration can be described by the spatially varying composition but that we do not have to specify the details of the molecular conformations. This can be justified if the time scale on which the composition changes is much larger than the molecular relaxation time, such that the chain conformations are always in "local equilibrium" with the instantaneous composition. The time scale for the chain conformations to equilibrate is set by $\tau = R_e^2/D$, where D denotes the single chain self-diffusion constant. Moreover, we assume that the free energy obtained from the SCF theory also provides an accurate description even out of equilibrium, i.e., not close to a saddle point.

Then, one can relate the spatial variation of the chemical potential to a current of the composition.

$$\boldsymbol{J}_\phi(\boldsymbol{r},t) = -\int \mathrm{d}^d\boldsymbol{r}'\, \Lambda_\phi(\boldsymbol{r},\boldsymbol{r}')\nabla_{\boldsymbol{r}'}\mu_\phi(\boldsymbol{r}',t) \tag{110}$$

$\boldsymbol{J}_\phi(\boldsymbol{r},t)$ denotes the current density at position $\boldsymbol{r}$ at time t. There is only one independent chemical potential, μ_ϕ, because of the incompressibility constraint and it is given by $\mu_\phi = \frac{\delta \mathcal{F}^*}{\delta \phi_A}$ with $\mathcal{F}^*[\Phi_A]$ from Eq. 53.

$$\frac{\mu_\phi(\boldsymbol{r})R_e^d}{\sqrt{\mathcal{N}}k_B T} = \chi N(\phi_B(\boldsymbol{r}) - \phi_A(\boldsymbol{r})) - (\mathrm{i}\Omega_A^*[\phi_A(\boldsymbol{r})] - \mathrm{i}\Omega_B^*[\phi_B(\boldsymbol{r})]) \tag{111}$$

$$\overset{\mathrm{RPA}}{\approx} (\chi N - 2)(1 - 2\phi_A(\boldsymbol{r})) - \frac{2R_e^2}{9}\triangle\phi_A(\boldsymbol{r}) \tag{112}$$

where $\mathrm{i}\Omega_A^*$ and $\mathrm{i}\Omega_B^*$ are real fields. In the last equation we have used the explicit expression for the chemical potential within RPA (cf. Eq. 96) for a symmetric binary polymer blend.

The kinetic Onsager coefficient, $\Lambda_\phi(\boldsymbol{r},\boldsymbol{r}')$, relates the "chemical force" acting on a monomer at position $\boldsymbol{r}'$ due to the gradient of the chemical potential to the concomitant current density at position $\boldsymbol{r}$. This describes purely relaxational dynamics, inertial effects are not captured. Λ_ϕ can be modelled in

different ways: The simplest approach would be a local coupling which results in the Onsager coefficient

$$\frac{\sqrt{\bar{\mathcal{N}}}k_{\mathrm{B}}T\Lambda_{\mathrm{local}}(\boldsymbol{r},\boldsymbol{r}')}{R_{\mathrm{e}}^2} = \frac{D}{R_{\mathrm{e}}^2}\phi_A(\boldsymbol{r}',t)\phi_B(\boldsymbol{r}',t)R_{\mathrm{e}}^d\delta(\boldsymbol{r}-\boldsymbol{r}') \quad \text{(local)} \tag{113}$$

The composition dependence accounts for the fact that currents of A- and B-densities have to exactly cancel in order to fulfill the incompressibility constraint. This local Onsager coefficient completely neglects the propagation of forces along the backbone of the chain and monomers move independently. Such a local Onsager coefficient is often used in calculations of dynamic models based on Ginzburg–Landau type energy functionals for reasons of simplicity [85–87].

Bearing in mind that the connectivity of monomeric units along the backbone of the polymer is an essential ingredient of the single chain dynamics it is clear that a non-local coupling should lead to a better description. In the Rouse model forces acting on a monomer caused by the chain connectivity are additionally taken into account [88, 89]. This leads to a kinetic factor that is proportional to the intramolecular pair-correlation function [90–92], $\mathcal{P}(\boldsymbol{r},\boldsymbol{r}')$

$$\frac{\sqrt{\bar{\mathcal{N}}}k_{\mathrm{B}}T\Lambda_{\mathrm{Rouse}}(\boldsymbol{r},\boldsymbol{r}')}{R_{\mathrm{e}}^2} = \frac{D}{R_{\mathrm{e}}^2}\phi_A(\boldsymbol{r}',t)\phi_B(\boldsymbol{r}',t)\mathcal{P}(\boldsymbol{r},\boldsymbol{r}') \qquad \text{(Rouse)}\,. \tag{114}$$

The difficulty in using Rouse dynamics lies in the computational expense of calculating the pair-correlation function in a spatially inhomogeneous environment at each time step. If the system is still fairly homogeneous (e.g., in the early stages of a demixing process), one can use the pair-correlation function of a homogeneous melt, as it is given through RPA [16]. This leads to the following Onsager coefficient in Fourier space:

$$\Lambda_{\mathrm{Rouse}}(q) = \Lambda_{\mathrm{local}}(q)g(q) = \frac{R_{\mathrm{e}}^{d+2}}{\sqrt{\bar{\mathcal{N}}}k_{\mathrm{B}}T}\frac{D}{R_{\mathrm{e}}^2}\tilde{\phi}_A\tilde{\phi}_B\frac{2(\exp[-x]-1+x)}{x^2} \tag{115}$$

x is defined as $x \equiv \boldsymbol{q}^2R_{\mathrm{e}}^2/6$ and g is the Debye function. Another model for non-local coupling is the reptation model [89, 93] which is appropriate for polymer melts with very long chains, i.e., entangled chains.

Since the total amount of A-component is conserved, the continuity equation relates the current of the composition to its time evolution:

$$\frac{\partial\phi_A(\boldsymbol{r},t)}{\partial t} + \nabla\boldsymbol{J}_\phi(\boldsymbol{r},t) = 0 \tag{116}$$

Eqs. 116 and 110 lead to the following diffusion equation:

$$\frac{\partial\phi_A(\boldsymbol{r},t)}{\partial t} = \nabla_{\boldsymbol{r}}\int \mathrm{d}^d\boldsymbol{r}'\,\Lambda_\phi(\boldsymbol{r},\boldsymbol{r}')\nabla_{\boldsymbol{r}'}\mu_\phi(\boldsymbol{r}',t) + \eta(\boldsymbol{r},t) \tag{117}$$

the last term representing noise that obeys the fluctuation-dissipation theorem. After Fourier transformation this diffusion equation adopts a very simple form:

$$\frac{\partial \phi_A(\boldsymbol{q},t)}{\partial t} = -\Lambda_\phi(\boldsymbol{q})q^2\mu_\phi(\boldsymbol{q},t) + \eta(\boldsymbol{q},t) \tag{118}$$

$$\overset{\text{RPA}}{\approx} -\frac{\sqrt{\mathcal{N}}k_\text{B}T\Lambda_\phi(qR_\text{e})^2}{R_\text{e}^{d+2}}\left[-2(\chi N-2)+\frac{2}{9}(qR_\text{e})^2\right]\phi_A+\eta \tag{119}$$

$$\overset{\text{RPA}}{\approx} -\frac{\sqrt{\mathcal{N}}k_\text{B}T\Lambda_\phi(qR_\text{e})^2}{R_\text{e}^{d+2}}\left[\frac{1}{\bar{\phi}_A g(q)}+\frac{1}{\bar{\phi}_B g(q)}-2\chi N\right]\phi_A+\eta \tag{120}$$

The last two equations are the explicit expressions within RPA, and $g(q)$ denotes the Debye function.

We have now found all necessary equations to numerically calculate the time evolution of the densities in a binary polymer mixture. This leads us to the following procedure which we refer to as the dynamic self consistent field theory (DSCFT) method:

1. Calculate the real fields $i\Omega_A^*$ and $i\Omega_B^*$ that "create" the density distribution according to Eq. 52 using the Newton–Broyden scheme.
2. Calculate the chemical potential μ_ϕ according to Eq. 111.
3. Propagate the density in time according to Eq. 118 using a simplified Runge–Kutta method.
4. Go back to (1).

5.2 External Potential Dynamics (EPD)

Instead of propagating the composition in time, we can study the time evolution of the exchange potential W. In equilibrium the density variable $\phi_A - \phi_B$ and the field variable $W = i\Omega_A - i\Omega_B$ are related to each other via $\phi_A - \phi_B = -W/\chi N$, see Eq. 25. We also use this identification to relate the time evolution of the field W to the time dependence of the composition. Since the composition is a conserved quantity and it is linearly related to the field variable, we also expect the field W with which we are now describing our system to be conserved. Therefore, we can use the free energy functional $\mathcal{H}[W]$ from Eq. 42 and describe the dynamics of the field W through the relaxational dynamics of a model B system, referring to the classification introduced by Hohenberg and Halperin [94].

$$\frac{\partial W(\boldsymbol{r},t)}{\partial t} = \nabla_r \int \mathrm{d}^d\boldsymbol{r}\Lambda_\text{EPD}(\boldsymbol{r},\boldsymbol{r}')\nabla_{r'}\mu_W(\boldsymbol{r}',t) + \eta(\boldsymbol{r},t) \tag{121}$$

with the chemical potential μ_W being the first derivative of the free energy $\mathcal{H}[W]$ with respect to the order parameter W,

$$\mu_W(\boldsymbol{r}) = \frac{\delta\mathcal{H}[W(\boldsymbol{r})]}{\delta W(\boldsymbol{r})} = \frac{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T}{R_\mathrm{e}^d}\frac{W + \chi N\left[\phi_A^*(\boldsymbol{r}) - \phi_B^*(\boldsymbol{r})\right]}{2\chi N}, \tag{122}$$

Λ_EPD is a kinetic Onsager coefficient, and η denotes noise that satisfies the fluctuation-dissipation theorem. The Fourier transform of this new diffusion equation is simply:

$$\frac{\partial W(q,t)}{\partial t} = -\frac{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T\Lambda_\mathrm{EPD}(qR_\mathrm{e})^2}{R_\mathrm{e}^{d+2}}\frac{W + \chi N(\phi_A^* - \phi_B^*)}{2\chi N} + \eta \tag{123}$$

$$\overset{\mathrm{RPA}}{\approx} -\frac{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T\Lambda_\mathrm{EPD}(qR_\mathrm{e})^2}{R_\mathrm{e}^{d+2}}\frac{W\left(1 - 2\chi N\bar{\phi}_A\bar{\phi}_B g(q)\right)}{2\chi N} + \eta \tag{124}$$

$\eta(q,t)$ is white noise that obeys the fluctuation-dissipation theorem. In the last equation we have used the Random-Phase Approximation for the Hamiltonian of the EP theory (cf. Eq. 90). We refer to the method which uses this diffusion equation as the external potential dynamics (EPD) [31]. As shown in Sect. 4.4.1, any Onsager coefficient (in conjunction with the fluctuation-dissipation theorem) will reproduce the correct thermodynamic equilibrium. But how is the dynamics of the field W related to the collective dynamics of composition fluctuations?

Comparing the diffusion equations of the dynamic SCF theory and the EP Dynamics, Eqs. 120 and 124, and using the relation $W(q,t) = -2\chi N\phi_A(q,t)$, we obtain a relation between the Onsager coefficients within RPA:

$$\Lambda_\mathrm{EPD}(q) = \frac{2\chi N}{\bar{\phi}_A\bar{\phi}_B g(q)}\Lambda_\phi(q) \tag{125}$$

In particular, the non-local Onsager coefficient Λ_Rouse that mimics Rouse-like dynamics (cf. Eq. 114) corresponds to a local Onsager coefficient in the EP Dynamics

$$\frac{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T\Lambda_\mathrm{EPD\text{-}Rouse}(q)}{R_\mathrm{e}^d} = 2\chi ND \tag{126}$$

Using the approximation,

$$\nabla_{\boldsymbol{r}}\mathcal{P}(\boldsymbol{r},\boldsymbol{r}') \cong -\nabla_{\boldsymbol{r}'}\mathcal{P}(\boldsymbol{r},\boldsymbol{r}') \tag{127}$$

one can derive Eq. 126 without invoking the random phase approximation. The approximation (Eq. 127) is obviously exactly valid for a homogeneous system, because the pair-correlation function only depends on the distance $|\boldsymbol{r} - \boldsymbol{r}'|$ between two points. It is expected to be reasonably good for weakly perturbed chain conformations [95].

Generally, one can approximately relate the time evolution of the field W to the dynamic SCF theory [31]. The saddle point approximation in the external

fields leads to a bijective relation between the external fields $i\Omega_A^*$, $i\Omega_B^*$ and ϕ_A, ϕ_B. In the vicinity of the saddle point, we can therefore choose with which of the two sets of variables we would like to describe the system.

To study the time evolution in the EPD we use the following scheme [29]:

0. Find initial fields, $W = i\Omega_A^* - i\Omega_B^*$ and $U = i\Omega_A^* + i\Omega_B^*$, that create the initial density distribution.
1. Calculate the chemical potential μ_W according to Eq. 122.
2. Propagate W according to Eq. 123 using a simplified Runge–Kutta method.
3. Adjust U to make sure the incompressibility constraint $\phi_A^* + \phi_B^* = 1$ is fulfilled again using the Newton–Broyden method.
4. Go back to (1).

The EPD method has two main advantages compared to DSCFT: First of all it incorporates non-local coupling corresponding to the Rouse dynamics via a local Onsager coefficient. Secondly it proves to be computationally faster by up to one order of magnitude. There are two main reasons for this huge increase in speed: In EPD the number of equations that have to be solved via the Newton–Broyden method to fulfill incompressibility is just the number of Fourier functions used. The number of equations in DSCFT that have to be solved to find the new fields after integrating the densities is twice as large. On the other hand, comparing the diffusion Eq. 117 used in the DSCFT method with Eq. 121 in EPD, it is easily seen, that the right hand side of the latter is a simple multiplication with the squared wave vector of the relevant mode, whereas the right hand side of Eq. 117 is a complicated multiplication of three spatially dependent variables.

6 Extensions

The dynamic self-consistent field theory has been widely used in the form of MESODYN [97]. This scheme has been extended to study the effect of shear on phase separation or microstructure formation, and to investigate the morphologies of block copolymers in thin films. In many practical applications, however, rather severe numerical approximations (e.g., very large discretization in space or contour length) have been invoked, that make a quantitative comparison to the the original model of the SCF theory difficult, and only the qualitative behavior could be captured.

More recently, stress and strain have been incorporated as a conjugated pair of slow dynamic variables to extend the model of the SCF theory. This allows us to capture some effects of viscoelasticity [76]. Similar to the evaluation of the single chain partition function by enumeration of explicit chain conformations, one can simulate an ensemble of mutually non-interacting chains exposed to the effective, self-consistent fields, U and W, in order to

obtain the densities [98]. Possibilities to extend this scheme to incorporate non-equilibrium chain conformations have been explored [96, 99].

7 Applications

7.1 Homopolymer-Copolymer Mixtures

As discussed in Sect. 4.1, one prominent example of a situation where the SCF theory fails on a qualitative level is the microemulsion channel in ternary mixtures of A and B homopolymers and AB diblock copolymers. Figure 5 shows an example of a mean-field phase diagram for such a system. Four different phases are found: A disordered phase, an ordered (lamellar) phase (see Fig. 4, right snapshot), an *A*-rich and a *B*-rich phase. The SCF theory predicts the existence of a point where all three phases meet and the distance of the lamellar sheets approaches infinity, an isotropic Lifshitz point [100, 101].

It seems plausible that fluctuations affect the Lifshitz point. If the lamellar distance is large enough that the interfaces between A and B sheets can bend around, the lamellae may rupture and form a globally disordered structure. A Ginzburg analysis reveals that the upper critical dimension of isotropic Lifshitz points is as high as 8 (see also Sect. 4.1). Unfortunately, the lower critical dimension of isotropic Lifshitz points is not known [102].

Indeed, the experimentally observed phase behavior differs substantially from the mean-field phase diagram. An example is shown in Fig. 6. The Lifshitz point is destroyed, the three phase coexistence region between the

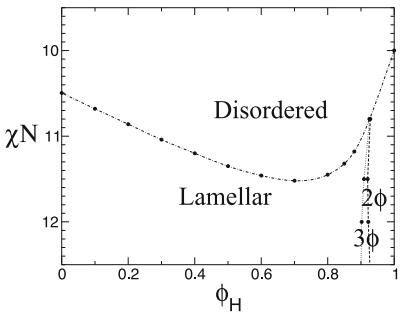

Fig. 5 Mean-field phase diagram for a ternary symmetric $A + B + AB$ blend as a function of the incompatibility parameter χN and the homopolymer volume fraction Φ_H. The chain lengths of the copolymers is five times that of the homopolymers. 2ϕ denotes a region of two-phase coexistence between an *A*-rich and a *B*-rich phase, 3ϕ one of three-phase coexistence between an *A*-rich, a *B*-rich, and a lamellar phase. From [80]

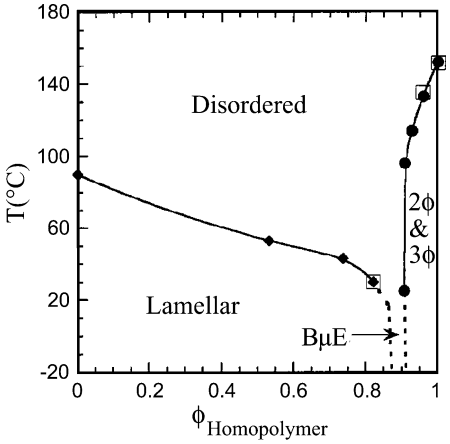

Fig. 6 Experimental phase diagram of a PDMS-PE/PDMS/PE blend with copolymers about five times as long as the homopolymers. From [103]. Reproduced by permission of the Royal Society of Chemistry

lamellar phase and the demixed *A*-rich and *B*-rich phases is removed and gives way to a macroscopically mixed phase. A transmission electron micrograph of this phase is shown and compared with the structure of the lamellar phase in Fig. 7.

In order to study this effect, Düchs et al. have performed field-theoretic Monte Carlo simulations of the system of Fig. 5 [80, 83, 104], in two dimensions. (For the reasons explained in Sect. 4.4.2, most of these simulations were carried out in the EP approximation). Some characteristic snapshots were already shown in Fig. 4. Here we show another series of snapshots at $\chi N = 12.5$ for increasing homopolymer concentrations (Fig. 8). For all these points, the self-consistent field theory would predict an ordered lamellar phase. In the

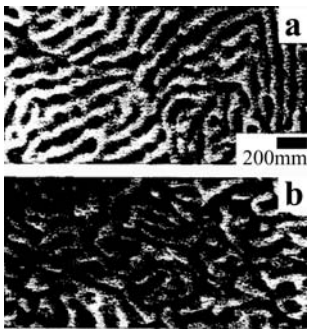

Fig. 7 Transmission electron micrographs from symmetric PE/PEP/PE-PEP blends: (**a**) fluctuating lamellae; (**b**) microemulsion phase. From [103]. Reproduced by permission of the Royal Society of Chemistry

Fig. 8 Snapshots of ϕ_A^* for the ternary $A + B + AB$ system of Fig. 5 at $\chi N = 12.5$ for homopolymer volume fraction $\phi_H = 0.74$, 0.78, 0.8, 0.82 (first row from left), and 0.84, 0.86, and 0.9 (second row from left)

simulation, only configurations with lower homopolymer concentrations exhibit (defective) lamellar order. At higher homopolymer concentration, the order breaks up and a microemulsion emerges.

The phase transition can be identified by defining appropriate order parameters. We use anisotropy parameters, which are defined in terms of the Fourier transform $F(\boldsymbol{q})$ of images as those shown in Fig. 8:

$$F_n(\boldsymbol{q}) \equiv \frac{1}{2\pi} \left| \int_0^{2\pi} \mathrm{d}\phi |F(\boldsymbol{q})|^2 \, \mathrm{e}^{\mathrm{i}n\phi} \right| . \tag{128}$$

$F_n(q)$ is zero for isotropic (disordered) configurations and non-zero for anisotropic (lamellar) configurations. The anisotropy information can be condensed into a single dimensionless parameter

$$\bar{F}_n := \frac{\int \mathrm{d}\boldsymbol{q} F_n(\boldsymbol{q})}{\sigma(\boldsymbol{q})|_{F_n}} , \tag{129}$$

with the normalization

$$\sigma(q)|_{F_n} \equiv \left[\frac{\int \mathrm{d}q \, q^2 F_n(q)}{\int \mathrm{d}q \, F_n(q)} - \left(\frac{\int \mathrm{d}q \, q F_n(q)}{\int \mathrm{d}q \, F_n(q)} \right)^2 \right]^{1/2} . \tag{130}$$

Examples of anisotropy parameters are shown in Fig. 9 (left). They are clearly suited to characterize the phase transition between the disordered phase at low χN and the lamellar phase at high χN. For comparison, the same system was also examined with complex Langevin simulations. The results are shown in Fig. 9 (right). The transition points are the same.

Unfortunately, the order of the transition could not yet be determined from these simulations. In the self-consistent field theory, the transition is continuous. According to theoretical predictions [54, 55], fluctuations shift it to lower χN and change the order of the transition. Our simulation data are

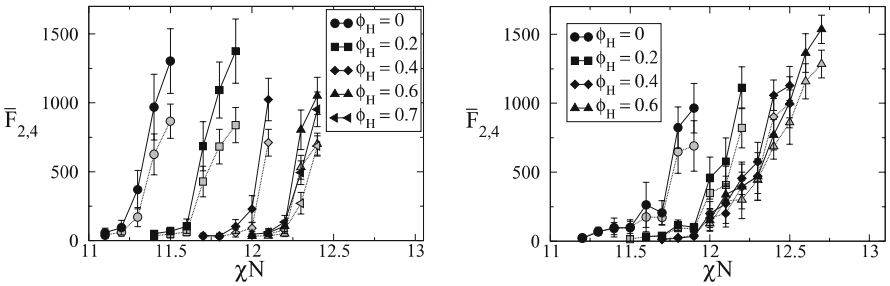

Fig. 9 Anisotropy parameters $\tilde{\mathcal{F}}_2$ (*solid*) and $\tilde{\mathcal{F}}_4$ (*dashed*) vs. χN for different homopolymer volume fractions Φ_H at the order/disorder transition. *Left*: From Monte Carlo Simulations (EP theory) *Right*: From Complex Langevin simulations. After [80]

consistent with the assumption that the transition is continuous or weakly first order. In particular, we have established that the miscibility gap between the ordered and the disordered phase is very small (data not shown).

Besides the order/disorder transition, we also need to determine the demixing transition. This was done in the grand canonical ensemble by monitoring the difference between A and B monomer densities. It exhibits a jump at the phase transition. All phase transition points are summarized in the phase diagram of Fig. 10. It is in good qualitative agreement with the experimental phase diagram of Fig. 6. In particular, we reproduce the cusp-like region of the microemulsion.

The simulations allow to analyze the phase transition and the microemulsion phase in more detail. Figure 11 compares different characteristic length scales of the system at the order/disorder transition. One length scale, L_0, is obtained from the maximum of $F_0(q)$ (Eq. 128) and gives the typical distance between layers. The other is the mean absolute curvature radius D_C,

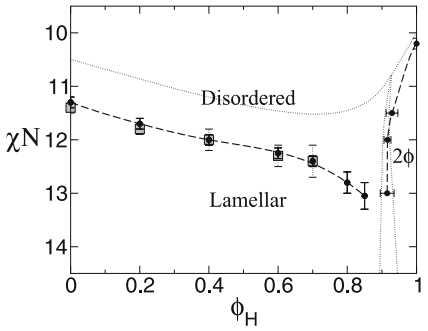

Fig. 10 Phase diagram from simulations for the same system as above. The *circles* show the results from Monte Carlo simulations, the *squares* those from the complex Langevin simulations. The *dotted lines* correspond to the mean-field prediction from Fig. 5. From [80]

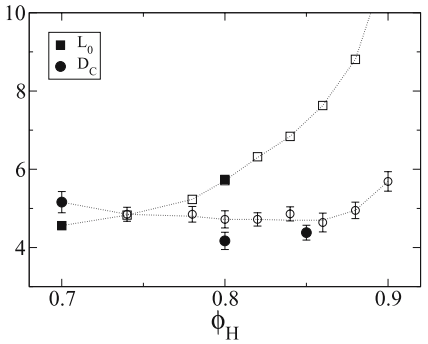

Fig. 11 Characteristic length scales at $\chi N = 12.5$ as a function of homopolymer concentration ϕ_H. D_C (*circles*) shows the mean (absolute) curvature radius, and L_0 the preferential length scale of layer distances, obtained from the maximum of $F_0(q)$ Eq. 128. From [80]

and characterizes the length scale for layer bending. Whereas the typical layer distance increases with the homopolymer concentration, the curvature radius remains roughly constant. The microemulsion phase transition takes place at the value of ϕ_H where the two lines cross. Thus we can identify the mechanism by which fluctuations generate the microemulsion: The lamellae break up when their width grows larger than the length scale which characterizes the boundary fluctuations.

To summarize, the example of homopolymer/copolymer mixtures demonstrates nicely how field-theoretic simulations can be used to study non-trivial fluctuation effects in polymer blends within the Gaussian chain model. The main advantage of these simulations is that they can be combined in a natural way with standard self-consistent field calculations. As mentioned earlier, the self-consistent field theory is one of the most powerful methods for the theoretical description of polymer blends, and it is often accurate on a quantitative level. In many regions of the parameter space, fluctuations are irrelevant for large chain lengths (large $\bar{\mathcal{N}}$) and simulations are not necessary. Field-theoretic simulations are well suited to complement self-consistent field theories in those parameter regions where fluctuation effects become important.

7.2 Spinodal Decomposition

To illustrate the application of the method to study the dynamics of collective composition fluctuations, we now consider the spontaneous phase separation (i.e., spinodal decomposition) that ensues after a quench into the miscibility gap at the critical composition, $\bar{\phi}_A = \bar{\phi}_B = 1/2$, of the symmetric blend. Different time regimes can be distinguished: During the early stages, composition fluctuations are amplified. Fourier modes with a wavevector q smaller

than a critical value q_c exponentially grow and modes with different wavevectors are decoupled. As the amplitude of the modes continues to grow, the composition in the domains begins to saturate and different modes begin to interact. At even later stages of phase ordering, hydrodynamics dominates the coarsening.

In the following, we restrict our attention to the early stages of spinodal decomposition. In the analysis of experiments one often uses the Landau–de Gennes functional (Eq. 96) which results in the Cahn–Hilliard–Cook theory [105] for the early stages of phase separation. This treatment predicts that Fourier modes of the composition independently evolve and increase exponentially in time with a wavevector-dependent rate, $\phi_A(q,t)^2 \sim \exp[R(\boldsymbol{q})t]$. Therefore, it is beneficial to expand the spatial dependence of the composition in our dynamic SCF or EP calculations in a Fourier basis of plane waves. As the linearized theory suggests a decoupling of the Fourier modes at early stages, we can describe our system by a rather small number of Fourier modes.

Using Eq. 120 one reads off the rate $R(q)$:

$$R(\boldsymbol{q}) = 4\frac{\sqrt{\bar{\mathcal{N}}}k_\mathrm{B}T\Lambda_\phi(qR_e^2)}{R_\mathrm{e}^{d+2}}\left[\chi N - 2 - \frac{(qR_\mathrm{e})^2}{9}\right] \tag{131}$$

and obtains for local dynamics and Rouse-like dynamics

$$\frac{R(\boldsymbol{q})R_\mathrm{e}^2}{D} = (qR_\mathrm{e})^2(\chi N - 2)\left[1 - \left(\frac{q}{q_\mathrm{c}}\right)^2\right] \tag{132}$$

$$\frac{R(\boldsymbol{q})R_\mathrm{e}^2}{D} = (qR_\mathrm{e})^2(\chi N - 2)g(q)\left[1 - \left(\frac{q}{q_\mathrm{c}}\right)^2\right], \tag{133}$$

respectively, where the critical wavevector is given by $q_\mathrm{c}R = \sqrt{9(\chi N - 2)}$. During the early stage, composition modes with wavevectors $q < q_\mathrm{c}$ will grow. In the case of local dynamics, the fastest growth occurs at $q_\mathrm{c}/\sqrt{2}$. In the case of Rouse-like dynamics, the maximum of the growth rate is shifted to smaller wavevectors. The expressions above use the RPA which becomes accurate at weak segregation (WSL). Of course, the applicability of the linearized theory of spinodal decomposition using the Landau–de Gennes free energy functional is severely limited: In the ultimate vicinity of the critical point, the free energy functional does not even give an accurate description of the thermodynamic equilibrium, and the relaxation times are characterized by a non-classical, dynamic critical exponent.

In the opposite limit of strong segregation (SSL) $\chi N \gg 2$, the free energy can also be expressed in terms of a square gradient functional, and one obtains $q_\mathrm{c}R = \sqrt{6(\chi N - 2)}$. At intermediate segregation, which is most relevant to experiments, the square-gradient approximation breaks down and there is

no straightforward way to interpolate between the simple analytical expression in the weak and strong segregation limit.

The predictions of the rate $R(q)$ for the weak and the strong segregation limit are presented in Fig. 12. As expected the results of the dynamic SCF theory with a local Onsager coefficient fall between the results of the Cahn–Hilliard–Cook theory in the two limits of weak and strong segregation. Comparing the predictions for the kinetics of phase separation to experiments or simulations, it often remains unclear whether deviations are caused by approximating the free energy functional by the Landau–de Gennes expressions or the Onsager coefficient. Indeed, earlier Monte Carlo simulations found rather pronounced deviations [106] from the prediction of the Cahn–Hilliard–Cook theory [105].

To overcome this difficulty we use EP dynamics or dynamic SCF theory, which do not invoke a gradient expansion of the free energy functional and yield an accurate description of the free energy costs of composition fluctuations. Quantitatively comparing dynamic calculations to computer simulations of the bond fluctuation model [107], we investigate the relation between the dynamics of collective composition fluctuations and the underlaying dynamics of single chains, which is encoded in the Onsager coefficient Λ.

The parameters of the bond fluctuation model can be related to the coarse-grained parameters, χN, R_e and $\bar{N}$, of the standard SCF model and the static properties of the bond fluctuation model (e.g., the bulk phase behavior [45, 53] and interface properties [44, 46, 57, 58]) have been quantitatively compared to SCF theory. Note that the mapping between the particle-based simulation model and the SCF model does not involve any adjustable parameter: the Flory–Huggins parameter χN is identified via the energy of mixing, the length scale is set by R_e and the time scale is determined by the self-diffusion constant of a single polymer chain in a dense melt. All these quantities are readily measurable in the simulations. In the following example, we

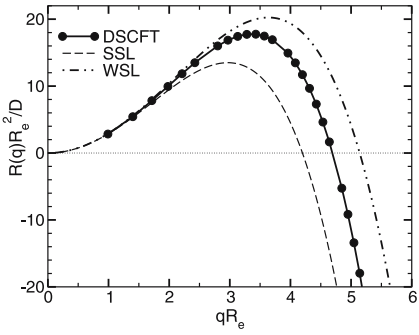

Fig. 12 Growth rates of composition modes for a quench of a symmetric mixture to $\chi N = 5$. Below the critical wave vector q_c the density modes increase spontaneously. Modes with larger wave vectors are damped. From [29]

consider a quench to $\chi N = 5$. This is a compromise: On the one hand it is sufficiently below the critical point, $\chi_c N = 2$, such that critical composition fluctuations are not important. On the other hand the interfaces between the coexisting phases are still wide enough such that the polymers are still describable by the Gaussian chain model on the length scale of the (intrinsic) interfacial width [108]. The chains in the bond fluctuation model comprise $N = 64$ segments which corresponds to an invariant degree of polymerization, $\bar{\mathcal{N}} = 240$. The time scale is set by $\tau = R_e^2/D = 1.5 \times 10^7$ Monte Carlo steps, where each segment attempts on average one random, local displacement during a Monte Carlo step. In the simulations we use a cubic system of size $L = 6.35R_e$ and average the results over 64 independent realizations of the temperature quench.

Figure 13 compares the time evolution of the dynamic structure factor as calculated by dynamic SCF theory using a local Onsager coefficient (panel a) and non-local Onsager coefficient (panel b) with the results of the Monte Carlo simulation. As qualitatively expected from the Cahn–Hilliard–Cook theory [105], there is a well-defined peak in the structure factor $S(\boldsymbol{q}, t)$

$$S(q,t) \equiv \left\langle \left| \int \mathrm{d}^d \boldsymbol{r} \, \left(\hat{\phi}_A(\boldsymbol{r},t) - \hat{\phi}_B(\boldsymbol{r},t) \right) \exp(\mathrm{i}\boldsymbol{q}\boldsymbol{r}) \right|^2 \right\rangle \tag{134}$$

that exponentially grows with time. For these calculations we have omitted the thermal noise. Therefore, composition fluctuations with $q < q_c$ grow, while those with $q > q_c$ are damped, and the structure factors for different times exhibit a common intersection at q_c. No such well-defined critical wavevector q_c is observed in Monte Carlo simulations or experiments,

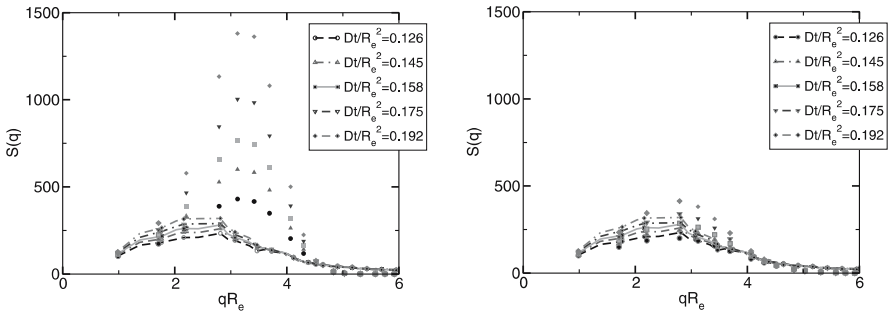

Fig. 13 Global structure factor versus wave vector for different times for a quench from the mixed state at $\chi N = 0.314$ to $\chi N = 5$. *Lines* represent Monte Carlo results, *symbols* dynamic SCF theory results. (**a**) Compares dynamic SCF theory using a local Onsager coefficient with Monte Carlo simulations. Local dynamics obviously overestimates the growth rate and shifts the wavevector that corresponds to maximal growth rate to larger values. (**b**) Compares dynamic SCF theory using a non-local Onsager coefficient that mimics Rouses dynamics with Monte Carlo results showing better agreement. From [29]

where thermal fluctuations limit the decay of fluctuations with $q > q_c$. Clearly, there is a sizeable difference between dynamic SCF calculations using a local Onsager coefficient and those which use a non-local Λ corresponding to Rouse-like dynamics. The peak of the structure factor in the dynamic SCF calculations with the local Onsager coefficient grows much too fast, and also the peak in the calculations occurs at larger wavevectors compared to the Monte Carlo simulations. Apparently, the calculations with the non-local Onsager coefficient agree much better with the simulation results.

To quantify this observation, we investigate the growth (or decay) of density modes. The data in Fig. 14 show that the time evolution during the early stages of phase separation indeed is exponential and one can define a growth rate. The results for a local Onsager coefficient are presented in Fig. 12 and are bracketed by the approximations for the weak and strong segregation limits.

In Fig. 15 we corroborate that the results of the dynamic SCF calculations using the non-local Onsager coefficient (Eq. 114) agree with the EPD using the local Onsager coefficient (Eq. 126) for the field W. As detailed in Sect. 5, however, the EPD calculations are computationally much less demanding.

The growth rates for the dynamic SCF theory with local Onsager coefficient are displayed in Fig. 16a and compared to the results of the Monte Carlo simulations. The maximal growth rate occurs at too large wavevectors and the growth rate for $q < q_c$ is too large. The Cahn plot, $R(q)/Dq^2$ vs. q^2 presented in the inset, shows a linear behavior (cf. Eq. 132) in contrast to the simulations.

Much of the discrepancy between calculations and Monte Carlo simulations is removed when we employ a non-local Onsager coefficient or EPD. For $q < q_c$ we obtain almost quantitative agreement for the growth rate at early times. In the Monte Carlo simulations, however, deviations from the expo-

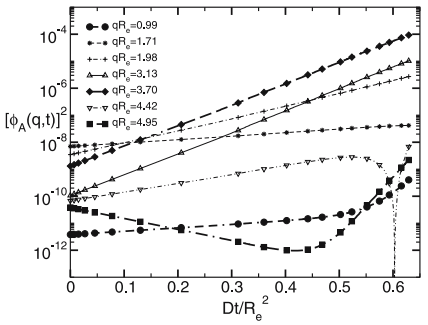

Fig. 14 Several density modes displayed on a logarithmic scale versus time. The results were obtained through DSCFT calculations in a three-dimensional system of length $L_x = L_y = L_z = 6.35R_e$ using $7 \times 7 \times 7$ functions for a quench from $\chi N = 0.314$ to $\chi N = 5$. The expected exponential behavior during early stages of demixing is well reproduced. From [29]

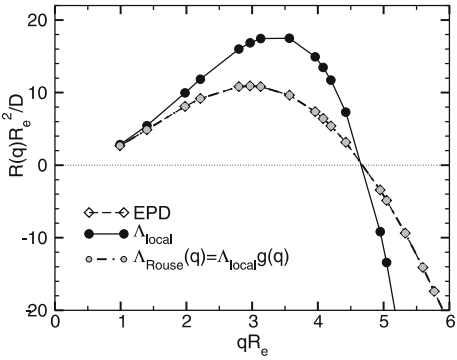

Fig. 15 Growth rates obtained through two dimensional DSCFT calculations using local and Rouse dynamics and EPD calculations. The DSCFT results using the pair-correlation function of a homogeneous melt and the EPD results are in good agreement. From [29]

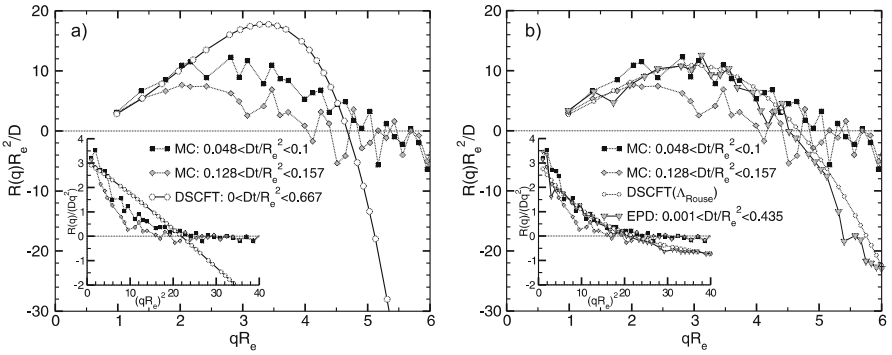

Fig. 16 Corresponding relaxation rates to Fig. 13. Panel (**a**) compares the Monte Carlo relaxation rates with DSCFT calculations with local dynamics. Panel (**b**) compares Monte Carlo results with EPD and DSCFT calculations with Rouse dynamics. For earlier times good agreement in the growth rate is found but Monte Carlo simulations show early deviations from the exponential behavior. From [29]

nential growth are observed earlier than in the calculations and there is no decay of fluctuations for $q > q_c$ in the simulations.

Both effects can be captured if we include fluctuations in the EPD calculations. The results of EPD simulations including fluctuations are presented in Fig. 17. For $q < q_c$ and early times the calculations with and without fluctuations mutually agree, and resemble the simulation results. At later times, the growth rate decreases in the EPD and the simulations. For $q > q_c$, fluctuations do not decay in the EPD but rather adopt their finite equilibrium strength. The growth rates extracted from the EPD for $q > q_c$ rather mirror the fact that some modes grow and others shrink during the time window of the simulations, although the results have been averaged over 64 independent, two-dimensional EPD calculations.

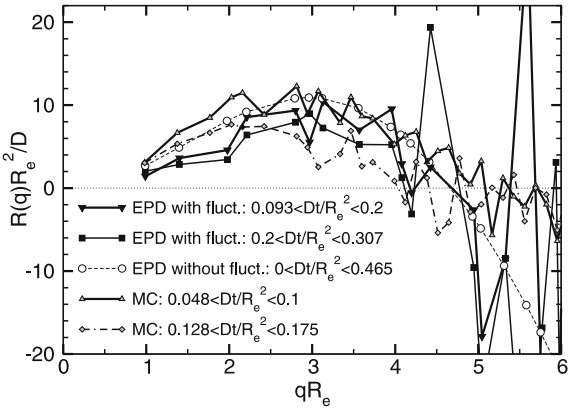

Fig. 17 Relaxation rates obtained through Monte Carlo simulations and EPD calculations with random fluctuations for different time intervals. Fluctuations lead in both methods to an earlier change in the exponential behavior of the growing modes. From [29]

The influence of the chain connectivity on the dynamics of the composition fluctuations does not only influence two-point correlation functions like the global structure factor but it is also visible in the time evolution of composition profiles in the vicinity of a surface [109].

The example highlights that both an accurate free energy functional as well as an Onsager coefficient that describes the dynamics of the molecules are important for a quantitative description. Of course, this study has concentrated on the simplest possible scenario. The assumption of the chain conformations being in equilibrium with the instantaneous composition field and the neglect of hydrodynamics have to be addressed to explore the behavior of more complex systems. Such a description of the kinetics of phase separation is important because many multi-component polymer systems do not reach equilibrium on large length scales and the morphology of interfaces or self-assembled structures depend on processing conditions.

8 Summary and Outlook

We have briefly reviewed methods which extend the self-consistent mean-field theory in order to investigate the statics and dynamics of collective composition fluctuations in polymer blends. Within the standard model of the self-consistent field theory, the blend is described as an ensemble of Gaussian threads of extension R_e. There are two types of interactions: zero-ranged repulsions between threads of different species with strength χN and an incompressibility constraint for the local density.

In the EP theory one regards fluctuation of the composition but still invokes a mean-field approximation for fluctuations of the total density. This approach is accurate for dense mixtures of long molecules, because composition fluctuations decouple from fluctuations of the density. The energy per monomer due to composition fluctuations is typically on the order of $\chi k_B T \sim k_B T/N$, and it is therefore much smaller than the energy of repulsive interactions (on the order of $1 k_B T$) in the polymer fluid that give rise to the incompressibility constraint. If there were a coupling between composition and density fluctuations, a better description of the repulsive hard-core interactions in the compressible mixture would be required in the first place.

Both the statics and the collective dynamics of composition fluctuations can be described by these methods, and one can expect these schemes to capture the essential features of fluctuation effects of the field theoretical model for dense polymer blends. The pronounced effects of composition fluctuations have been illustrated by studying the formation of a microemulsion [80]. Other situations where composition fluctuations are very important and where we expect that these methods can make straightforward contributions to our understanding are, e.g., critical points of the demixing in a polymer blend, where one observes a crossover from mean field to Ising critical behavior [51, 52], or random copolymers, where a fluctuation-induced microemulsion is observed [65] instead of macrophase separation which is predicted by mean-field theory [64].

The application of the dynamic SCF theory [97] or EPD [29, 31, 109] to the collective dynamics of concentration fluctuations and the relation between the dynamics of collective concentration fluctuations and the single chain dynamics is an additional, practically important aspect. We have merely illustrated the simplest possible case—the early stages of spontaneous phase separation within purely diffusive dynamics. In applications the hydrodynamic effects [110, 111], shear and viscoelasticity [112] might become important. Even deceptively simple situations—like nucleation phenomena in binary polymer blends—still pose challenging questions [113]. Also the assumption of local equilibrium for the chain conformations, which allows us to use the SCF free energy functional, has to be questioned critically. Methods have been devised to incorporate some of these complications [76, 96, 99, 111, 112] but the development in this area is still in its early stages.

Field theoretical simulations [74, 75, 80] avoid any saddle point approximation and provide a formally exact solution of the standard model of the self-consistent field theory. To this end one has to deal with a complex free energy functional as a function of the composition and density. This significantly increases the computational complexity. Moreover, for certain parameter regions, it is very difficult to obtain reliable results due to the sign problem that a complex weight imparts onto thermodynamical averages [80]. We have illustrated that for a dense binary blend the results of the field theoretical simulations and the EP theory agree quantitatively, i.e., density and composi-

tion fluctuation decouple [80]. In other systems, however, density fluctuations are of crucial importance. Let us mention two examples: (i) Field theoretical simulation methods can be used to study semi-dilute solutions [75, 77], i.e., a compressible ensemble of Gaussian threads with repulsive interactions. In dilute solutions the chains adopt a self-avoiding walk behavior, while in semi-dilute solutions the excluded volume interactions are screened on a large length scale [16]. The very same model has been extensively studied by series expansions [114] and renormalization group theory [115] and those results provide an excellent testing bed to gauge the efficiency of the field theoretical simulation method. (ii) It is a challenge to apply field theoretical simulations to particle-based models. This might provide valuable insight into the relation between coarse-grained parameters, R_e and χN, of the standard model of the self-consistent field theory and the parameters (temperature, pressure/density and chain length) of a particle-based model.

Acknowledgements It is a great pleasure to thank K. Binder, D. Düchs, G.H. Fredrickson, V. Ganesan, E. Reister, and M. Schick for their enjoyable and fruitful collaboration on this topic. Financial support was provided by the DFG under grants Mu 1674/1, Mu 1674/3, and Schm 985/7 and the calculations were performed at the NIC in Jülich and the HLR Stuttgart.

References

1. Cahn RW, Haasen P, Kramer EJ (1993) Materials Science and Technology, A Comprehensive Treatment, Vol. 12. Wiley-VCH, Weinheim; Garbassi F, Morra M, Occhiello E (2000) Polymer Surface: From Physics to Technology. Wiley, Chichester, 510 pp
2. Shull KR, Mayes AM, Russell TP (1993) Macromolecules 26:3929; Mayes AM, Russell TP, Satija SK, Majkrzak CF (1992) Macromolecules 25:6523; Russell TP, Anastasiadis SH, Menelle A, Flecher GP, Satija SK (1991) Macromolecules 24:1575; Green PF, Russell TP (1991) Macromolecules 24:2931
3. Dai KH, Norton LJ, Kramer EJ (1994) Macromolecules 27:1949; Dai KH, Kramer EJ (1994) Polymer 35:157; Dai KH, Kramer EJ, Shull KR (1992) Macromolecules 25:220; Jones RAL, Kramer EJ, Rafailovich MH, Schwarz SA (1989) Phys Rev Lett 62:280
4. Budkowski A, Klein J, Fetters L (1995) Macromolecules 28:8571
5. Tanaka H, Hasegawa H, Hashimoto T (1991) Macromolecules 24:240
6. Jansen BJP, Rastogi S, Meijer HEH, Lemstra PJ (2001) Macromolecules 34:3998
7. Tucker CL, Moldenaers P (2002) Ann Rev Fluid Mechanics 34:177
8. Edwards SF (1965) Proc Phys Soc 85:613
9. Helfand E, Tagami Y (1971) J Polym Sci B 9:741, J Chem Phys 56:3592; 57:1812 (1972); Helfand E, Sapse AM (1975) J Chem Phys 62:1327; Helfand E (1975) J Chem Phys 62:999
10. Hong KM, Noolandi J (1981) Macromolecules 14:727; 14:737
11. Scheutjens JMHM, Fleer GJ (1979) J Chem Phys 83:1619
12. Matsen MW, Schick M (1994) Phys Rev Lett 72:2660; Matsen MW (1995) Phys Rev Lett 74:4225; (1995) Macromolecules 28:5765; (2001) J Phys: Cond Matt 14:R21
13. Schmid F (1998) J Phys: Cond Matt 10:8105

14. Flory PJ (1941) J Chem Phys 9:660; Huggins HL (1941) J Chem Phys 9:440
15. In $d = 2$ dimensions $\bar{\mathcal{N}}$ is independent from the number of segments N per molecule and mean field theory is inaccurate, cf. Cavallo A, Müller M, Binder K (2003) Europhys Lett 61:214
16. de Gennes PG (1979) Scaling Concepts in Polymer Physics. Cornell University Press, Ithaca, 319 pp
17. Wittmer JP et al. (2004) Phys Rev Lett 93:147601
18. Kratky O, Porod G (1949) Rec Trav Chim 68:1106; Saito N, Takahashi K, Yunoki Y (1967) J Phys J Soc Jpn 22:219
19. Szleifer I (1997) Curr Opin Colloid Interface Sci 2:416; Szleifer I, Carignano MA (1996) Adv Chem Phys 94:742
20. Müller M, Schick M (1996) Macromolecules 29:8900
21. Müller M, Mac LGDowell (2000) Macromolecules 33:3902; Müller M, Mac LGDowell, Yethiraj A (2003) J Chem Phys 118:2929
22. Müller M (1998) Macromolecules 31:9044
23. Matsen MW (2002) J Phys: Cond Matt 14:21
24. Rowlinson JS, Swinton FL (1982) Liquids and liquid mixtures. Butterworths, London; Van Konynenburg P, Scott RL (1980) Philos Trans Soc London Series A298:495
25. Barker A, Henderson D (1967) J Chem Phys 47:4714
26. Hansen JP, Donald Mc IR (1986) Theory of simple liquids. Academic Press, New York, 568 pp
27. Orwoll RA, Arnold PA (1996) Polymer-Solvent Interaction Parameter χ. In: Mark JE (ed) Physical Properties of Polymers, Chapter 14. AIP, Woodbury, New York
28. Reister E (2001) PhD Thesis, Johannes Gutenberg-Universität, Mainz, Germany
29. Reister E, Müller M, Binder K (2001) Phys Rev E 64:041804
30. We note that on the right hand side of Eq. 32 the second term is smaller than the first term by a factor of R_e^d/V and can therefore be neglected in large systems. $R_e^d \delta\phi_\alpha^*(\boldsymbol{r})/\delta W_\beta(\boldsymbol{r}')$ is independent of the system size
31. Maurits NM, Fraaije JGEM (1997) J Chem Phys 107:5879
32. Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1986) Numerical Recipes. Cambridge Univ Press, Cambridge
33. Yamamoto T (2000) J Comp App Math 124:1; Chen X (1997) J Comp App Math 80:105; Martinez JM (2000) J Comp App Math 124:45
34. Schmid F, Müller M (1995) Macromolecules 28:8639
35. Anderson DG (1965) J Assoc Comput Mach 12:547
36. Eyert V (1996) J Comp Phys 124:271
37. Thompson RB, Rasmussen Kø, Lookman T (2004) J Chem Phys 120:31
38. Drolet F, Frederickson GH (1999) Phys Rev Lett 83:4317
39. Mitchell AR, Griffiths DF (1980) The Finite Difference Method in Partial Differential Equations. Wiley, Chichester, 284 pp
40. Feit MD, Fleck JA Jr, Steiger A (1982) J Comp Phys 47:412
41. Rasmussen Kø, Kalosakas G (2002) J Polym Sci B 40:777
42. Düchs D (2003) PhD Thesis, Univ of Bielefeld, Germany
43. Morse DC, Fredrickson GH (1994) Phys Rev Lett 73:3235
44. Schmid F, Müller M (1995) Macromolecules 28:8639
45. Müller M (1999) Macromol Theory Simul 8:343
46. Müller M, Binder K (1998) Macromolecules 31:8323
47. Amundson K, Helfand E, Davis DD, Quan X, Patel SS (1991) Macromolecules 24:6546; Lin CY, Schick M, Andelman D (2005) Macromolecules 38:5766
48. Li XJ, Schick M (2001) J Biophys 80:1703

49. Ginzburg VL (1960) Sov Phys Solid State 1:1824; de Gennes PG (1977) J Phys Lett (Paris) 38:L-441; Joanny JF (1978) J Phys A 11:L-117; Binder K (1984) Phys Rev A 29:341
50. Fisher ME (1974) Rev Mod Phys 46:587
51. Gehlen MD, Rosedale JH, Bates FS, Wignall GD, Almdal K (1992) Phys Rev Lett 68:2452; Schwahn D, Meier G, Mortensen K, Janssen S (1994) J Phys II (France) 4:837; Frielinghaus H, Schwahn D, Willner L, Springer T (1998) Physica B 241:1022
52. Deutsch HP, Binder K (1992) Macromolecules 25:6214; (1993) J Phys II (France) 3:1049
53. Müller M, Binder K (1995) Macromolecules 28:1825
54. Brazovskii SA (1975) Sov Phys JETP 41:85
55. Fredrickson GH, Helfand E (1987) J Chem Phys 87:697
56. Buff FP, Lovett RA, Stillinger FH (1965) Phys Rev Lett 15:621
57. Werner A, Schmid F, Müller M, Binder K (1997) J Chem Phys 107:8175
58. Werner A, Schmid F, Müller M, Binder K (1999) Phys Rev E 59:728
59. Müller M, Albano EV, Binder K (2000) Phys Rev E 62:5281
60. Müller M, Schick M (1996) J Chem Phys 105:8885
61. Holyst R, Schick M (1992) J Chem Phys 96:7728
62. Taupin C, de Gennes PG (1982) J Phys Chem 86:2294
63. Müller M, Gompper G (2002) Phys Rev E 66:041805
64. Shakhnovich EI, Gutin AM (1989) J Phys France 50:1843; Fredrickson GH, Milner ST, Leibler L (1992) Macromolecules 25:6341; Fredrickson GH, Milner ST (1991) Phys Rev Lett 67:835; Subbotin AV, Semenov AN (2002) Eur Phys J 7:49
65. Houdayer J, Müller M (2002) Europhys Lett 58:660; Houdayer J, Müller M (2004) Macromolecules 37:4283
66. Shi AC, Noolandi J, Desai RC (1996) Macromolecules 29:6487
67. Laradji M, Shi AC, Noolandi J, Desai RC (1997) Phys Rev Lett 78:2577; (1997) Macromolecules 30:3242
68. Montvay I, Münster G (1994) Quantum fields on the lattice. Cambridge Univ Press, Cambridge, 505 pp
69. Klauder JR (1983) J Phys A 16:L317
70. Parisi G (1983) Phys Lett 131B:393
71. Lin HQ, Hirsch JE (1986) Phys Rev B 34:1964
72. Baeurle SA (2002) Phys Rev Lett 89:080602
73. Moreira AG, Baeurle SA, Fredrickson GH (2003) Phys Rev Lett 91:150201
74. Ganesan V, Fredrickson GH (2001) Europhys Lett 55:814
75. Fredrickson GH, Ganesan V, Drolet F (2002) Macromolecules 35:16
76. Fredrickson GH (2002) J Chem Phys 117:6810
77. Alexander-Katz A, Moreira AG, Fredrickson GH (2003) J Chem Phys 118:9030
78. Gausterer H (1998) Nucl Phys 642:239
79. Schoenmakers WJ (1987) Phys Rev D 36:1859
80. Düchs D, Ganesan V, Fredrickson GH, Schmid F (2003) Macromolecules 36:9237
81. Yethiraj A, Schweizer KS (1993) J Chem Phys 98:9080; Schweizer KS, Yethiraj A (1993) J Chem Phys 98:9053
82. Holyst R, Vilgis TA (1993) J Chem Phys 99:4835; (1994) Phys Rev E 50:2087
83. Düchs D, Schmid F (2004) NIC Series 20:343
84. Duane S, Kennedy AD, Penleton BJ, Roweth D (1987) Phys Lett B 195:216; Mehlig B, Heermann DW, Forrest BM (1992) Phys Rev B 45:679; Forrest BM, Suter UW (1994) J Chem Phys 101:2616
85. Kotnis M, Muthukumar M (1992) Macromolecules 25:1716

86. Chakrabarti A, Toral R, Gunton JD, Muthukumar M (1989) Phys Rev Lett 63:2072
87. Castellano C, Glotzer SC (1995) J Chem Phys 103:9363
88. Rouse PE (1953) J Chem Phys 21:1272
89. Doi M, Edwards SF (1994) The Theory of Polymer Dynamics. Oxford University Press, Oxford, 402 pp
90. Binder K (1983) J Chem Phys 79:6387
91. de Gennes PG (1980) J Chem Phys 72:4756
92. Pincus P (1981) J Chem Phys 75:1996
93. de Gennes PG (1971) J Chem Phys 55:572
94. Hohenberg PC, Halperin BI (1977) Rev Mod Phys 49:435
95. Particle-based SCF schemes like "single chain in mean field"–simulations [96] avoid the use of an Onsager coefficient and are applicable to systems with strong spatial inhomogeneities
96. Müller M, Smith GD (2005) J Polym Sci B 43:934
97. Altevogt P, Evers OA, Fraaije JGEM, Maurits NM, van Vlimmeren BAC (1999) J Mol Struc Theochem 463:139; Fraaije JGEM (1993) J Chem Phys 99:9202
98. Soga KG, Zuckermann MJ, Guo H (1996) Macromolecules 29:1998; Besold G et al. (2000) J Polym Sci 38:1053
99. Ganesan V, Pryamitsyn V (2003) J Chem Phys 118:4345
100. Hornreich RM, Luban M, Shtrikman S (1975) Phys Rev Lett 35:1678
101. Hornreich RM (1980) J Magn Magn Mat 15:387
102. Diehl HW (2002) Acta Physica Slovaca 52:271
103. Morkved TL, Chapman BR, Bates FS, Lodge TP, Stepanek P, Almdal K (1999) Faraday Discuss 11:335
104. Düchs D, Schmid F (2004) J Chem Phys 121:2798
105. Cahn JW, Hilliard JE (1958) J Chem Phys 28:258; (1959) ibid 31:668; Cook HE (1970) Acta Metall 18:297
106. Sariban A, Binder K (1989) Polym Comm 30:205; (1991) Macromolecules 24:578; Baumgärtner A, Heermann DW (1986) Polymer 27:1777
107. Carmesin I, Kremer K (1988) Macromolecules 21:2819; Wittmann HP, Kremer K (1990) Comp Phys Com 61:309; Deutsch HP, Binder K, (1991) J Chem Phys 94:294
108. Müller M, Werner A (1997) J Chem Phys 107:10–764
109. Reister E, Müller M (2003) J Chem Phys 118:8476
110. Maurits NM, Zvelindovsky AV, Fraaije JGEM (1998) J Chem Phys 108:9150; Koga T, Kawasaki K, Takenaka M, Hashimoto T (1993) Physica A198:473
111. Keesta BJ, van Puyvelde PCJ, Anderson PD, Heijer HEH (2003) Phys Fluids 15:2567
112. Tanaka HM (2000) J Phys: Cond Mat 12:R207; (1999) Prog Theo Phys 101:863
113. Lefebvre AA, Lee JH, Balsara NP, Vaidyanathan C (2002) J Chem Phys 117:9063 and 9074
114. Muthukumar M, Nickel BG (1987) J Chem Phys 86:460
115. des Cloizeaux J, Jannink G (1990) Polymers in Solution: Their Modeling and Structure, Oxford Science Publications, Oxford; Freed KF (1987) Renormalization Group Theory of Macromolecules. Wiley, New York; Le Guillou JC, Zinn-Justin J (1989) J Phys (France) 50:1365; Schäfer L (1999) Excluded Volume Effects in Polymer Solutions. Springer, Berlin, Heidelberg, New York

Adv Polym Sci (2005) 185: 59–109
DOI 10.1007/b136793

Published online: 30 August 2005

Efficient Methods to Compute Long-Range Interactions for Soft Matter Systems

Axel Arnold[1,2] (✉) · Christian Holm[1,2]

[1]Max-Planck-Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Germany
arnolda@mpip-mainz.mpg.de, *holm@mpip-mainz.mpg.de*

[2]Frankfurt Institute for Advanced Studies (FIAS), Johann Wolfgang Goethe-Universität, Frankfurt/Main, Germany
A.Arnold@fias.uni-frankfurt.de, *c.holm@fias.uni-frankfurt.de*

Abstract An extensive introduction to the topic of how to compute long-range interactions efficiently is presented. First, the traditional Ewald sum for 3D Coulomb systems is reviewed, then the P^3M method of Hockney and Eastwood is discussed in some detail, and alternative ways of dealing with the Coulomb sum are briefly mentioned. The best strategies to perform the sum under partially periodic boundary conditions are discussed, and two recently developed methods are presented, namely the MMM2D and ELC methods for two-dimensionally periodic boundary conditions, and the MMM1D method

for systems with only one periodic coordinate. The dipolar Ewald sum is also reviewed. For some of the methods, error formulas are provided which enable the algorithm to be tuned at a predefined accuracy.

Keywords Ewald methods · Long-range interactions · Coulomb sum · Periodic boundary conditions · Dipolar sum

1 Introduction

Soft condensed matter (or soft matter, as it is often called) is a term for materials in states of matter that are neither simple liquids nor hard solids of the type studied, for example, in solid-state physics. Many such materials are familiar from everyday life—glues, paints, soaps, baby diapers—while others are important in industrial processes, such as polymer melts that are molded and extruded to form plastics [1]. Biological materials are mainly made out of soft matter as well: membranes, actin filaments, DNA, RNA, and proteins belong to this class. Furthermore, most of the food we digest is soft matter. All these materials share the importance of having length scales intermediate between the atomic and macroscopic scales: the relevant range for soft matter lies between nanometers and micrometers. Examples are polymers, colloids, liquid crystals, glasses, and dipolar fluids. Typical energies between different structures are similar to thermal energies. Hence, Brownian motion or thermal fluctuations play a prominent role. Another key feature of soft matter systems is their propensity to self-assemble. The energy differences during this process are small such that many neighboring states are normally accessible through thermal fluctuations. This often results in complex phase behaviors yielding a rich variety of accessible structures. Order does not necessarily arise only on the single-molecule level, but quite commonly exhibits a multitude of hierarchically ordered structures of sometimes tremendous intricacy and complexity. Most of the biological systems are usually not even in equilibrium but evolve among switchable steady states.

Given this wide field, research on soft material substances often acquires knowledge from different areas, such as physics, chemistry, and biology. Therefore, a high level of interdisciplinary cooperation may be required for certain scientific questions.

In the past, our research has mainly focused on the study of charged polymers (polyelectrolytes) and charged colloids, which serve as important substances for many technical applications. Charged systems also occur in biological environments (since most biological matter is charged) [2], and modeling of explicit water molecules requires partial charges as well.

Computational simulations provide some unique ways to elucidate the properties of charged systems. However, they are time-consuming, since the

interaction between charged species is long ranged. Moreover, to compute the interactions of all particles (here denoted by N) quickly and accurately, one needs complex algorithms to beat the unfavorable $\mathcal{O}(N^2)$ complexity which one obtains from simply counting all particle interactions.

In this contribution we give a short overview of state-of-the-art methods of treating Coulomb interactions in fully or partially periodic geometries, and also show how to treat dipolar interactions. We try to emphasize the strengths and weaknesses of various methods to facilitate the right choice of method for a particular problem. We highlight the level of control over the unavoidable systematic error that each method intrinsically possesses. At the end we give a short description of our simulation package ESPResSo [3, 4] that contains most of the presented methods. The source code of this software, and consequently of the methods, is freely available as source code from http://www.espresso.mpg.de.

2
Methods for the Coulomb Sum in Fully Periodic Boundary Conditions

For many physical investigations one wants to simulate bulk properties and therefore introduce periodic boundary conditions in all spatial directions to avoid boundary effects. For these kinds of boundary conditions, the famous Ewald sum [5–8] does a remarkable job in splitting the very slowly converging sum over the Coulomb potential into two exponentially converging sums. Still, this method suffers from two deficiencies. First, it is computationally demanding. This is due to the fact that one part of the problem is solved in reciprocal space, implying the need for several Fourier transformations. Second, the algorithm scales like N^2 with N being the number of charged particles in the primary box, or at best like $N^{3/2}$, if one uses cutoffs which are optimized with respect to the splitting parameter [9]. The situation becomes worse if only partially periodic boundary conditions are applied.

In fully periodic boundary conditions, the Ewald method can be accelerated to a computation time scaling of $N \log N$ using fast Fourier transformation (FFT) methods by replacing the charges with a regular mesh. Various mesh-based methods exist, such as P^3M, PME, or SPME, but P^3M is known to be the computationally optimal variant [10].

A completely different approach to the problem is due to Lekner [11] and Sperb [12, 13]. Although the method in the original Lekner formulation has $\mathcal{O}(N^2)$ complexity, Strebel and Sperb have developed a factorization approach which yields an $\mathcal{O}(N \log N)$ algorithm [14], called MMM[1]. The Strebel approach can also be adapted for partially periodic systems [15]; however, the method becomes computationally more expensive.

[1] The acronym MMM is due to R. Strebel and has no obvious meaning.

Other advanced methods of order $\mathcal{O}(N \log N)$ are tree algorithms [16], which are the first-order approximation of even better so-called fast multipole methods [17] (FMM). These methods can reach a linear complexity, but at the expense of a heavy computational overhead, which makes them advantageous only for a very large number of charges ($N \approx 100\,000$) [18]. Another recent approach is to use multigrid methods which solve the Poisson equation in real space [19, 20]. These have the advantage that they should be more useful on massively parallel architectures, since an efficient implementation of 3D FFTs is difficult to achieve on those machines. An additional advantage could be the fact that for molecular dynamics (MD) simulations the old Poisson solution can still be used as a starting point for the multigrid iteration, thereby improving drastically the convergence speed. Unlike the mesh-based Ewald methods or the Sperb method, these methods can be adapted for partially periodic boundary conditions without increasing the computational complexity. However, these methods become much slower, and therefore the crossover to the other methods might only happen at very large values of N.

In this section we give an introduction to the Ewald summation, collecting the important equations for energy and forces. We also discuss briefly the fast multipole method and MMM. In the next section we present some recent results of our research on how to deal with partially periodic boundary conditions. Finally, we briefly discuss a new lattice method due to Maggs [21]. The material has been mainly collected from the sources [10, 15, 22–27]. As good textbooks for background material we recommend the second edition of Frenkel and Smit [28] and the book by Allen and Tildesley [29].

2.1 The Standard 3D Ewald Method

There are many examples of long-range interactions that can be treated by Ewald techniques [29]. Here, however, only the case of Coulomb point charges will be discussed, i.e., an interaction potential $1/r$. Consider therefore a system of N particles with charges q_i at positions $\boldsymbol{r}_i$ in an overall neutral and, for simplicity, cubic simulation box of length L. If periodic boundary conditions are applied, the total electrostatic energy of the box is given by

$$E = \frac{1}{2} \sum_{\boldsymbol{m} \in \mathbb{Z}^3}^{\infty} \sum_{i,j=1}^{N}{}' \frac{q_i q_j}{|\boldsymbol{r}_{ij} + \boldsymbol{m}L|}\,, \tag{1}$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$, $\boldsymbol{n}$ counts the periodic images, and the prime denotes that the summand for $i = j$ has to be omitted for $\boldsymbol{n} = 0$. Since this sum is only conditionally convergent, its value is not well-defined unless one specifies the precise way in which the cluster of simulation boxes should fill the $\mathbb{R}^3$, i.e., its *shape*, e.g., approximately spherical [7, 8, 30, 31]. Usually a spherical limit

is applied:

$$E = \frac{1}{2} \sum_{S=0}^{\infty} \sum_{\boldsymbol{m}^2=S} \sum_{i,j=1}^{N}{}' \frac{q_i q_j}{|\boldsymbol{r}_{ij} + \boldsymbol{m}L|} . \tag{2}$$

In Sect. 3.2 we will have to deal with a different order of summation. If one adds up the particles along z slabwise, i.e., ordered by increasing z-distance, but radially in x and y, Smith [32] has shown that the difference between the slabwise summed energy and the spherical summed up energy is given by

$$2\pi M_z^2 - \frac{2\pi \boldsymbol{M}^2}{3} , \tag{3}$$

where

$$\boldsymbol{M} = (M_x, M_y, M_z) = \sum_{i=1}^{N} q_i \boldsymbol{r}_i \tag{4}$$

is the net dipole moment of the primary simulation box. We will see below that the Ewald sum contains a summand $2\pi \boldsymbol{M}^2/3$, which is just exchanged by $2\pi M_z^2$ by this change of order.

The term $2\pi \boldsymbol{M}^2/3$ is called the dipole term. Besides determining the order of summation, it has another interesting property. The spherical summation order is equivalent to the limit of a large, spherically bounded, regular grid of images of the simulation box embedded in vacuum, basically a crystalline ball (although for many particles the "crystal" may be quite complex). If the surrounding space is filled by a homogeneous medium with a dielectric constant ε', the particles of the ball will feel a polarization force. It can be shown that this leads to an additional contribution that will not vanish even in the limit of an infinite ball, although then the complete space is filled by copies of the simulation box. The additional contribution for the infinite ball is again given by a modification of only the dipole term, which then reads

$$\frac{2\pi \boldsymbol{M}^2}{2\varepsilon' + 1} . \tag{5}$$

This again shows that even in the limit of a fully filled space the Coulomb sum "remembers" the way the summation was done, which is of course only possible due to its conditional convergence.

It is often stated that the long-range nature of the Coulomb potential complicates the treatment of electrostatic interactions. This is, however, only one part of the problem. In fact, the Coulomb potential bears *two* intrinsic difficulties. It is slowly decaying at large distances, and strongly varying at small distances. It is the combination of these two properties which leads to severe problems. If only one of them was present everything would be comparatively easy, since a short-range potential could be treated by a simple cutoff, as is

done, for example, for interactions of the Lennard–Jones type, and a long-range potential, which is periodic and slowly varying *everywhere*, can be accurately represented by the first few terms of its Fourier series.

Obviously, each of the two complications forbids the simple solution of the other, and the slowly decaying long-range part of the Coulomb potential renders a straightforward summation of Eq. 1 impracticable. The trick is thus to split the problem into two parts by the trivial identity

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1-f(r)}{r}. \tag{6}$$

The underlying idea is to distribute the two complications between the two terms in Eq. 6 by a suitable choice of the splitting function f. In particular:

- The first part $\frac{f(r)}{r}$ should be negligible, or even zero, beyond some cutoff $r_{\max}$, so that the summation up to the cutoff is a good approximation to (or the exact result of) this contribution to the total electrostatic potential.
- The second part $\frac{1-f(r)}{r}$ should be a slowly varying function for *all* r, so that its Fourier transform can be represented by only a few $\boldsymbol{k}$-vectors with $|\boldsymbol{k}| \leq k_{\max}$. This permits an efficient calculation of this contribution to the total electrostatic potential in reciprocal space.

Since the field equations are linear, the sum of these two contributions gives the solution for the potential of the original problem.

The two requirements on the splitting function f mentioned above leave a large freedom of choice [6, 33, 34]. The traditional selection is the complementary error function $\operatorname{erfc}(r) := \frac{2}{\sqrt{\pi}} \int_r^\infty \mathrm{d}t\, \mathrm{e}^{-t^2}$. This results in the well-known Ewald formula for the electrostatic energy of the primary box:

$$E = E^{(r)} + E^{(k)} + E^{(s)} + E^{(d)}, \tag{7}$$

where $E^{(r)}$ is the contribution from real space, $E^{(k)}$ the contribution from reciprocal space, $E^{(s)}$ the self-energy, and $E^{(d)}$ the dipole correction. They can be written as [28, 29]

$$E^{(r)} = \frac{1}{2} \sum_{\boldsymbol{m}\in\mathbb{Z}^3} {\sum_{i,j}}' q_i q_j \frac{\operatorname{erfc}(\alpha|\boldsymbol{r}_{ij}+\boldsymbol{m}L|)}{|\boldsymbol{r}_{ij}+\boldsymbol{m}L|} \tag{8}$$

$$E^{(k)} = \frac{1}{2}\frac{1}{V} \sum_{\boldsymbol{k}\neq 0} \frac{4\pi}{k^2} \mathrm{e}^{-k^2/4\alpha^2} |\tilde{\rho}(\boldsymbol{k})|^2 \tag{9}$$

$$E^{(s)} = -\frac{\alpha}{\sqrt{\pi}} \sum_i q_i^2 \tag{10}$$

$$E^{(d)} = \frac{2\pi}{(1+2\varepsilon')V} \left(\sum_i q_i \boldsymbol{r}_i \right)^2, \tag{11}$$

where the Fourier-transformed charge density $\tilde{\rho}(\boldsymbol{k})$ is defined as

$$\tilde{\rho}(\boldsymbol{k}) = \int_V \mathrm{d}^3 r \rho(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\boldsymbol{r}} = \sum_{j=1}^{N} q_j \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\boldsymbol{r}_j} \tag{12}$$

$$\text{where} \quad \boldsymbol{k} \in \frac{2\pi}{L}\mathbb{Z}^3 .$$

The advantage of rewriting Eq. 1 this way is that the exponentially converging sums over $\boldsymbol{m}$ and $\boldsymbol{k}$ in Eqs. 8 and 9 allow the introduction of comparatively small cutoffs without much loss in accuracy. Typically one chooses α large enough as to employ the minimum image convention in Eq. 8. The inverse length α, which is often referred to as the Ewald (or splitting) parameter, tunes the relative weights of the real-space and the reciprocal-space contributions. However, the final result of the exact Eq. 7, not terminating the sums at some finite cutoff value, is independent of α. The form of Eq. 10 given for the dipole correction assumes that the set of periodic replications of the simulation box tends spherically toward an infinite cluster, and that the medium outside this sphere is a homogeneous dielectric with dielectric constant ε' [7], whereas inside we set for simplicity $\varepsilon = 1$. The derivation of this term is not straightforward and requires an accurate mathematical treatment of the conditional convergence of Eq. 1, and of the image charges generated in the corresponding dielectric medium. Note that the case of a surrounding vacuum corresponds to $\varepsilon' = 1$ and that the dipole correction vanishes for metallic (or "tinfoil") boundary conditions, since then $\varepsilon' = \infty$. A detailed discussion of this term can be found in [7, 8, 31]. The dipole correction in Eq. 10 is independent of the Ewald parameter α. This again shows that the correction is not specific to the Ewald sum but more generally reflects the problems inherent to the conditional convergence of the $\boldsymbol{n}$ sum in Eq. 1. For the computation of $E^{(d)}$ the particle coordinates must *not* be folded back into the primary unit cell, for otherwise each boundary crossing produces an unphysical jump in the electrostatic energy (see [30]). The cutoffs $r_{\max}$ and $k_{\max}$ can be optimized with respect to α such that the required computer time scales like $N^{3/2}$ (see Perram et al. [9]). This, however, may require that $r_{\max} > L/2$ prohibiting the simple minimum image convention in real space and rendering this procedure less tempting. For given finite real- and reciprocal-space cutoffs there exists an optimal α value such that the accuracy of the approximated Ewald sum is the highest possible. This optimal value can be determined easily with the help of the excellent estimates for the cutoff errors derived by Kolafa and Perram [35], essentially by demanding that the real- and reciprocal-space contribution to the error should be equal. If the system under investigation is not electrostatically neutral, the infinite sum in Eq. 1 diverges. It can be made convergent by adding a homogeneously distributed background charge which restores neutrality—a typical situation for one-component plasma simulations. This results in an additional electroneutrality term $E^{(n)}$ to be included

in Eq. 7, which reads (see, e.g., [36])

$$E^{(n)} = -\frac{\pi}{2\alpha^2 V}\left(\sum_i q_i\right)^2 . \tag{13}$$

Since the neutralizing background is homogeneous, the correction term in Eq. 13 is independent of the particle positions.

The force $\boldsymbol{F}_i$ on particle i is obtained by differentiating the electrostatic potential energy E with respect to $\boldsymbol{r}_i$, i.e.,

$$\boldsymbol{F}_i = -\frac{\partial}{\partial \boldsymbol{r}_i} E . \tag{14}$$

Using Eqs. 7–12 one obtains the following Ewald formula for the forces:

$$\boldsymbol{F}_i = \boldsymbol{F}_i^{(r)} + \boldsymbol{F}_i^{(k)} + \boldsymbol{F}_i^{(d)} , \tag{15}$$

with the real-space, Fourier-space, and dipole contributions given by:

$$\begin{aligned}\boldsymbol{F}_i^{(r)} = q_i \sum_j q_j \sum_{\boldsymbol{m}\in\mathbb{Z}^3}{}' \Bigg(& \frac{2\alpha}{\sqrt{\pi}} \exp(-\alpha^2 |\boldsymbol{r}_{ij} + \boldsymbol{m}L|^2) \\ & + \frac{\operatorname{erfc}(\alpha|\boldsymbol{r}_{ij} + \boldsymbol{m}L|)}{|\boldsymbol{r}_{ij} + \boldsymbol{m}L|} \Bigg) \frac{\boldsymbol{r}_{ij} + \boldsymbol{m}L}{|\boldsymbol{r}_{ij} + \boldsymbol{m}L|^2}\end{aligned} \tag{16}$$

$$\boldsymbol{F}_i^{(k)} = q_i \sum_j q_j \frac{1}{V} \sum_{\boldsymbol{k}\neq 0} \frac{4\pi \boldsymbol{k}}{k^2} \exp\left(-\frac{k^2}{4\alpha^2}\right) \sin(\boldsymbol{k}\cdot\boldsymbol{r}_{ij}) \tag{17}$$

$$\boldsymbol{F}_i^{(d)} = -\frac{4\pi q_i}{(1+2\varepsilon')V} \sum_j q_j \boldsymbol{r}_j . \tag{18}$$

Since the self-energy in Eq. 10 and the neutralizing contribution in Eq. 13 are independent of particle positions, they do not contribute to the force.

2.2 The Particle Mesh Ewald Idea

The Fourier transformations involved in Eqs. 9 and 17 are the most time-consuming part of the Ewald sum. Several methods have been proposed to address this problem, e.g., tabulation of the complete Ewald potential [37] or the use of polynomial approximations. A particularly successful approach for the latter is the expansion of the nonspherical contributions to the Ewald potential in cubic harmonics [37, 38]. Apart from the difficulty of computational overheads which may strongly increase with the desired accuracy, these methods do not solve the additional problem of unfavorable scaling with particle number. At best they scale as $N^{3/2}$, which is more costly than a plain cutoff or reaction field approach.

The essential idea is *not* to avoid the Fourier transformations but to modify the problem in such a way that it permits application of the FFT (see, e.g., [39] and references therein). This reduces the complexity of the reciprocal part of the Ewald sum to $N \log N$. For an increasing number of particles the real-space cutoff can be varied in such a way that this scaling applies to the complete Ewald sum. Since the FFT is a grid transformation, there are discretization problems to be solved and corresponding discretization errors to be minimized. Performing the Fourier transformations in the reciprocal space part of the Ewald sum by FFT routines is by no means straightforward:

1. The point charges with continuous coordinates have to be replaced by a grid based charge density, since the FFT is a discrete and finite transformation.
2. It is neither obvious nor true that the best grid approximation to the continuum solution of the Poisson equation is achieved by using the continuum Green function [40].
3. There are at least three possibilities for implementing the differentiation needed in Eq. 14. They differ in accuracy and speed.
4. The procedure of assigning the forces calculated on the mesh back to the actual particles can, under certain circumstances, lead to unwanted violations of Newton's third law. They can be anything between harmless and disastrous.

The four steps involved in a particle mesh calculation are sources of various kinds of errors originating, for example, from discretization, interpolation, or aliasing[2] problems. Since these contributions are not independent of each other (reducing one might enhance another), the only reasonable demand is the minimization of the *total* error at a given computational effort.

The original literature on particle mesh routines is usually not easy to digest for the nonspecialist. It is obscured by the fact that various authors approach the problem from different directions and use different notations. There exist three mesh implementations of the Ewald sum—similar in spirit but different in detail. The oldest is the original particle–particle–particle mesh (P^3M) method of Hockney and Eastwood [40], and then there are two variants, namely, the particle mesh Ewald (PME) method of Darden et al. [41] and an extension of the latter by Essmann et al. [42], which is usually referred to as the smooth particle mesh Ewald (SPME) method.

In two papers by Deserno et al. [10, 22] it was shown how the three methods differ in detail, and it was demonstrated that the oldest method, namely the original P^3M algorithm, is actually the most accurate one. Since in addition error estimates exist [22], this mesh method should be the preferred method of choice. For historical reasons, however, many program packages

[2] A finite grid cannot represent arbitrarily large $\boldsymbol{k}$-vectors. Instead, they are folded back into the first "Brillouin zone" and distort there the true spectrum. This effect is usually referred to as "aliasing". See, e.g., Sect. 12.1 in [39].

have implemented the SPME method. We still have some hope that this will change in the future. The i$\boldsymbol{k}$ differentiation of PME and the analytic differentiation of the charge assignment function in real space of SPME can both also be used in P^{3}M, making it the most versatile mesh algorithm on the market. Our experience shows that the break-even point between a standard Ewald sum and the P^{3}M algorithm is around 600–800 charges, depending on the implementation details and desired accuracy. The algorithm can also be efficiently parallelized [43, 44].

2.3 P^{3}M in a Nutshell

The P^{3}M method maps the system onto a mesh, such that the necessary Fourier transformations can be accomplished by fast Fourier routines. At the same time the simple Coulomb Green function $4\pi/k^2$ is adjusted to make the result of the mesh calculation most closely resemble the continuum solution.

The first step, i.e., generating the mesh-based charge density ρ_{M} (defined at the mesh points $\boldsymbol{r}_p$), is carried out with the help of a charge assignment function W:

$$\rho_{\mathrm{M}}(\boldsymbol{r}_p) = \frac{1}{h^3} \sum_{i=1}^{N} q_i W(\boldsymbol{r}_p - \boldsymbol{r}_i)\,. \tag{19}$$

Here h is the mesh spacing, and the number of mesh points $N_{\mathrm{M}} = L/h$ along each direction should preferably be a power of two, since in this case the FFT is most efficient. The charge assignment function is classified according to its order P, i.e., between how many grid points per coordinate direction each charge is distributed. For W a cardinal B-spline [45] is chosen, which is a piecewise polynomial function of weight one. The order P gives the number of sections in the function. The first seven cardinal B-splines are sketched in Fig. 1. Its Fourier transform is

$$\widetilde{W}(\boldsymbol{k}) = h^3 \left(\frac{\sin(\frac{1}{2}k_x h)}{\frac{1}{2}k_x h} \frac{\sin(\frac{1}{2}k_y h)}{\frac{1}{2}k_y h} \frac{\sin(\frac{1}{2}k_z h)}{\frac{1}{2}k_z h} \right)^P . \tag{20}$$

In a second step the mesh-based electric field $\boldsymbol{E}(\boldsymbol{r}_p)$ is calculated. Basically, the electric field is the derivative of the electrostatic potential, but there exist several alternatives for implementing the differentiation on a lattice [10]. Here we will restrict ourselves to the case of i$\boldsymbol{k}$ differentiation, which works by multiplying the Fourier-transformed potential with i$\boldsymbol{k}$. In this case $\boldsymbol{E}(\boldsymbol{r}_p)$ can be written as

$$\boldsymbol{E}(\boldsymbol{r}_p) = \overleftarrow{\mathrm{FFT}} \left[-\mathrm{i}\boldsymbol{k} \cdot \overrightarrow{\mathrm{FFT}}\,[\rho_{\mathrm{M}}] \cdot \widehat{G}_{\mathrm{opt}} \right] (\boldsymbol{r}_p)\,. \tag{21}$$

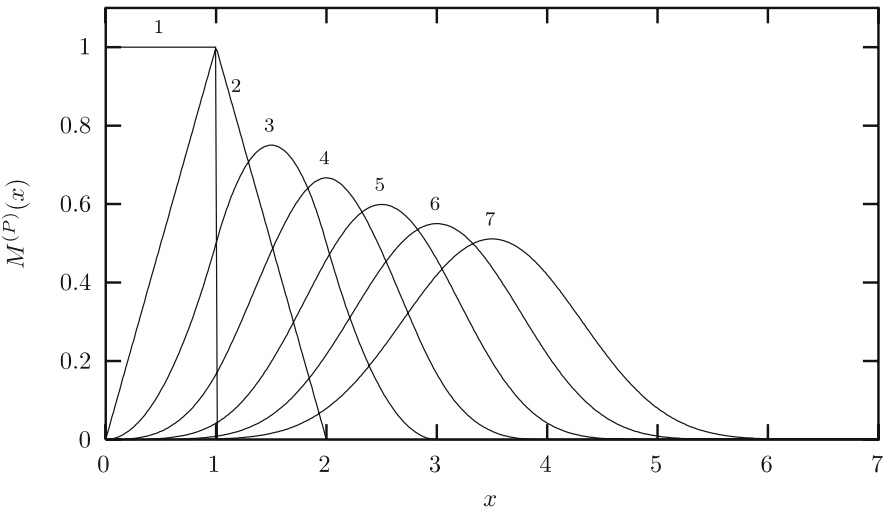

Fig. 1 Sketch of the first seven cardinal B-splines $M^{(P)}(x)$, parameterized by P. Note that the charge assignment functions $W^{(P)}(x)$ for the P^3M algorithm are just the "centered" B-splines

In words, $\boldsymbol{E}(\boldsymbol{r}_p)$ is the *backward* finite Fourier transform of the product of $-\mathrm{i}\boldsymbol{k}$, the *forward* finite Fourier transform of the mesh-based charge density ρ_{M}, and the so-called optimal influence function $\hat{G}_{\mathrm{opt}}$, given by

$$\widehat{G}_{\mathrm{opt}}(\boldsymbol{k}) = \frac{\widetilde{\boldsymbol{D}}(\boldsymbol{k}) \cdot \sum_{\boldsymbol{m}\in\mathbb{Z}^3} \widetilde{U}^2(\boldsymbol{k} + \frac{2\pi}{h}\boldsymbol{m}) \widetilde{\boldsymbol{R}}(\boldsymbol{k} + \frac{2\pi}{h}\boldsymbol{m})}{|\widetilde{\boldsymbol{D}}(\boldsymbol{k})|^2 \left[\sum_{\boldsymbol{m}\in\mathbb{Z}^3} \widetilde{U}^2(\boldsymbol{k} + \frac{2\pi}{h}\boldsymbol{m})\right]^2} \tag{22}$$

$$\text{with} \qquad \widetilde{\boldsymbol{R}}(\boldsymbol{k}) := -\mathrm{i}\boldsymbol{k}\frac{4\pi}{k^2}\,\mathrm{e}^{-k^2/4\alpha^2} \tag{23}$$

$$\text{and} \qquad \widetilde{U}(\boldsymbol{k}) := \widetilde{W}(\boldsymbol{k})/h^3\,. \tag{24}$$

Here $\widetilde{\boldsymbol{D}}(\boldsymbol{k})$ is the Fourier transform of the employed differentiation operator, which is simply $\mathrm{i}\boldsymbol{k}$ in our case. Finally, one arrives at the force on particle i, i.e., the replacement of Eq. 17:

$$\boldsymbol{F}_i = q_i \sum_{\boldsymbol{r}_p\in\mathbb{M}} \boldsymbol{E}(\boldsymbol{r}_p) W(\boldsymbol{r}_i - \boldsymbol{r}_p)\,. \tag{25}$$

Hereby, the sum extends over the complete mesh $\mathbb{M}$.

Although the presented formulas in Eqs. 19–25 look somewhat complicated, they are rather easy to implement in a step-by-step procedure. If the real-space cutoff $r_{\max}$ is chosen small enough (so that the real-space contribution Eq. 16 can be calculated in order N), the complete algorithm is essentially of order $N \log N$ (see the article by Petersen [46] for a nice exposition of this point).

2.4 How and Why to Control Errors

An investigation of the errors connected with particle-mesh Ewald methods is important for three reasons:

1. The procedure of discretization introduces new sources of errors in addition to the ones originating from real- and reciprocal-space cutoffs.
2. Comparing the efficiency of different mesh methods is only fair if it is done at the same level of accuracy.
3. The tuning parameters should be chosen in such a way as to run the algorithm at its optimal operation point.

The two last points, of course, apply to *any* numerical method to compute energy or forces for the Coulomb or dipolar sum. Error estimates can tell us if we might see artifacts in simulations due to too small cutoffs, or if our observations have some other origin. They can tell us how the algorithm scales at its optimal point and they can help us save a lot of expensive computer time. Unfortunately, some of the free or commercially available computer programs choose parameter combinations automatically, according to some more or less known rules. Often, the user is not aware of the applied approximations, which is a very dangerous route, since, after all, one needs to interpret the data. Therefore, we stress here the point that for all our implemented routines where we have error estimates, we make use of them.

However, there is no unique or optimal measure of accuracy. If MD simulations are performed, the main interest lies in errors connected with the force, while in Monte Carlo (MC) simulations one is concerned with accurate energies. In the simulation of ensemble averages it is the global accuracy, measured, e.g., by root-mean-square quantities, which is important. In the simulation of rare events local accuracy and maximal errors are relevant as well. Errors in the force can be due to their magnitude or due to their direction. And finally, one might be interested in absolute or relative errors.

Whatever quantity one decides to look at, it can be investigated as a function of system parameters like particle separation or distribution, *tuning* parameters like α, mesh size, or interpolation order and *components* of the algorithm, e.g., interpolation or differentiation scheme or splitting function $f(r)$. This obviously gives rise to a very large number of combinations. In other words, the corresponding parameter space is large and nontrivial. That is, general statements concerning the performance of one method can usually not be extracted from low-dimensional cuts through this space, because different methods scale differently with respect to their parameters.

We are concerned with one measure of accuracy, namely the root-mean-square (rms) error in the force given by

$$\Delta F := \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{F}_i - \boldsymbol{F}_i^{\mathrm{exa}}\right)^2}\,, \tag{26}$$

where $\boldsymbol{F}_i$ is the force on particle number i calculated via some mesh method and $\boldsymbol{F}_i^{\mathrm{exa}}$ is the *exact* physical force on that particle, calculable, e.g., by a well-converged standard Ewald sum. There exist error estimates for the real-space and Fourier-space contribution to this error for the standard Ewald sum [35], for the PME method [46], and for the P^3M method [22].

Here we show the dependence of the rms error in the force in Eq. 26 on the number of charged particles and their valence[3]. Since the assumptions and arguments involved are of a rather general nature, the result is not specific to a certain type of Ewald method.

It is reasonable to assume that the error in the force on particle i can be written as

$$\Delta\boldsymbol{F}_i := \boldsymbol{F}_i - \boldsymbol{F}_i^{\mathrm{exa}} = q_i \sum_{j\neq i} q_j \boldsymbol{\chi}_{ij}\,. \tag{27}$$

The idea behind this ansatz is that—just as it is true for $\boldsymbol{F}_i$—the *error* in $\boldsymbol{F}_i$ originates from the $N-1$ interactions of particle i with the other charged particles, and each contribution should be proportional to the product of the two charges involved. The vector $\boldsymbol{\chi}_{ij}$ gives the direction and magnitude of this error for two unit charges, and depends on their separation and orientation as well as on the specific algorithm used for calculating the electrostatic forces. For this term it is further assumed that

$$\left\langle \boldsymbol{\chi}_{ij}\cdot\boldsymbol{\chi}_{ik}\right\rangle = \delta_{jk}\left\langle \boldsymbol{\chi}_{ij}^2\right\rangle =: \delta_{jk}\chi^2\,, \tag{28}$$

where averaging over the particle configurations is denoted by the angular brackets $\langle\cdots\rangle$. The underlying assumption that contributions from different particles are uncorrelated is certainly not always true; think, for example, of highly ordered or strongly inhomogeneous particle distributions. However, it is a sensible one for *random* systems. Obviously, the term $\langle\boldsymbol{\chi}_{ij}^2\rangle$—the mean-square force error for two unit charges—can no longer depend on i and j, and is thus written as χ^2. Using Eqs. 27 and 28, it follows that

$$29\left\langle(\Delta\boldsymbol{F}_i)^2\right\rangle = q_i^2\sum_{j\neq i}\sum_{k\neq i} q_j q_k\langle\boldsymbol{\chi}_{ij}\cdot\boldsymbol{\chi}_{ik}\rangle \approx q_i^2\chi^2\mathcal{Q}^2\,, \tag{29}$$

[3] Note that this is by no means the only interesting measure of accuracy.

where the important quantity $\mathcal{Q}^2$ is defined as

$$\mathcal{Q}^2 := \sum_{j=1}^{N} q_j^2 \,. \tag{30}$$

We further assume that

$$\left\langle \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(\Delta \boldsymbol{F}_i\right)^2} \right\rangle \approx \sqrt{\frac{1}{N}\sum_{i=1}^{N} \langle \left(\Delta \boldsymbol{F}_i\right)^2 \rangle} \,, \tag{31}$$

which can be shown as true for reasonably large systems by the law of large numbers, along the line of reasoning in Ref. [22]. We end up with the final relation

$$\Delta F \approx \chi \frac{\mathcal{Q}^2}{\sqrt{N}} \,. \tag{32}$$

Thus, the scaling of the rms error in the force with particle number and valence is given by the factor $\mathcal{Q}^2 N^{-1/2}$, whereas the prefactor χ—which cannot be obtained by such simple arguments—contains the details of the method and is independent of the simulated system. Note that any information on the valence distribution enters only through the value of $\mathcal{Q}^2$.

The most interesting ingredient of the P^{3}M method is the optimal influence function from Eq. 22. It is constructed such that the result of the mesh calculation is as close as possible to the solution of the original continuum problem. More precisely, the P^{3}M method is derived from the requirement that the resulting Fourier-space contribution to the force minimizes the the following error measure Q:

$$Q := \frac{1}{h^3} \int_{h^3} \mathrm{d}^3 r_1 \int_{L^3} \mathrm{d}^3 r \left[\boldsymbol{F}(\boldsymbol{r}; \boldsymbol{r}_1) - \boldsymbol{R}(\boldsymbol{r}) \right]^2 \tag{33}$$

$\boldsymbol{F}(\boldsymbol{r}; \boldsymbol{r}_1)$ is the Fourier-space contribution of the force between two unit charges at positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_1 + \boldsymbol{r}$ as calculated by the P^{3}M method (note that due to broken rotational and translational symmetry this does in fact depend on the coordinates of *both* particles), and $\boldsymbol{R}(\boldsymbol{r})$ is the corresponding exact reference force (whose Fourier transform is just Eq. 23). The inner integral over $\boldsymbol{r}$ scans all particle separations, whereas the outer integral over $\boldsymbol{r}_1$ averages over all possible locations of the first particle within a mesh cell. Obviously, up to a factor L^{-3} this expression is just the mean-square error in the force for two unit charges, in other words, the quantity χ^2 from Eq. 28. This provides a link between the rms error of an N particle system and the error Q from Hockney and Eastwood. Using Eq. 32 one obtains

$$\Delta F \approx \mathcal{Q}^2 \sqrt{\frac{Q}{NL^3}} \,. \tag{34}$$

It is important to realize that Hockney and Eastwood not only provide a closed expression for the optimal influence function $\widehat{G}_{\text{opt}}$, but also a closed expression for the corresponding "optimal error" $Q_{\text{opt}} = Q[\widehat{G}_{\text{opt}}]$:

$$Q_{\text{opt}} = \frac{1}{L^3} \sum_{\boldsymbol{k} \in \widehat{\mathbb{M}}} \left\{ \sum_{\boldsymbol{m} \in \mathbb{Z}^3} \left| \widetilde{\boldsymbol{R}}(\boldsymbol{k} + \frac{2\pi}{h}\boldsymbol{m}) \right|^2 - \frac{\left| \widetilde{\boldsymbol{D}}(\boldsymbol{k}) \sum_{\boldsymbol{m} \in \mathbb{Z}^3} \widetilde{U}^2(\boldsymbol{k} + \frac{2\pi}{h}\boldsymbol{m}) \widetilde{\boldsymbol{R}}^*(\boldsymbol{k} + \frac{2\pi}{h}\boldsymbol{m}) \right|^2}{|\widetilde{\boldsymbol{D}}(\boldsymbol{k})|^2 \left[\sum_{\boldsymbol{m} \in \mathbb{Z}^3} \widetilde{U}^2(\boldsymbol{k} + \frac{2\pi}{h}\boldsymbol{m}) \right]^2} \right\} . \tag{35}$$

The outer sum extends over all $\boldsymbol{k}$-vectors of the Fourier-transformed mesh $\widehat{\mathbb{M}}$, and the asterisk denotes the complex conjugate. Once again, in the special case of i$\boldsymbol{k}$ differentiation one has $\widetilde{\boldsymbol{D}}(\boldsymbol{k}) = \mathrm{i}\boldsymbol{k}$.

Admittedly, Eq. 35 looks rather complicated. Still, in combination with Eq. 34 it gives the rms force error of the $\mathrm{P^3M}$ method or, more precisely, of its Fourier-space contribution. After all, the computation of Q_{opt} and that of $\widehat{G}_{\text{opt}}$ are quite similar. It should be emphasized that the formula Eq. 35 for the optimal Q value, just like the one for the optimal influence function in Eq. 22, is of a very general nature. It also works for different charge assignment functions, reference forces, or any differentiation scheme which can be expressed by an operator $\widetilde{\boldsymbol{D}}(\boldsymbol{k})$.

The corresponding rms error in the force from the real-space contribution in Eq. 16 has been derived by Kolafa and Perram [35] and is provided here for reference purposes:

$$\Delta F^{(r)} \approx \frac{2\mathcal{Q}^2}{\sqrt{N r_{\max} L^3}} \exp(-\alpha^2 r_{\max}^2) . \tag{36}$$

With these two estimates at hand it is easy to determine the optimal value of the splitting parameter α via a stand-alone program, which takes the relevant system parameters (N, $\mathcal{Q}^2$, L) and specifications of the algorithm ($r_{\max}$, N_{M}, P) as its input. If real- and reciprocal-space contributions to the error, $\Delta F^{(r)}$ and $\Delta F^{(k)}$, respectively, are assumed to be statistically independent, the total error is given by

$$\Delta F = \sqrt{\left(\Delta F^{(r)}\right)^2 + \left(\Delta F^{(k)}\right)^2} . \tag{37}$$

This quantity has to be minimized with respect to α. In most cases it is, however, accurate enough to use the following approximation: determine the value of α at which the real- and reciprocal-space contribution to the rms force error are equal.

2.5 FMM

The key to the improved scaling of the Ewald method is a product decomposition of the Fourier-space sum. In contrast, the multipole methods [17, 47, 48] are based on a product decomposition in real space. This requires a product decomposition of the plain $1/r$ potential. Now let $x = (r, \theta, \phi)$ and $y = (r', \theta', \phi')$ be the spherical coordinates of two points, $r' < r$, and let $\gamma = \angle(x, y)$ be the angle between the origin and x and y. Then $|x - y| = r^2 + r'^2 - 2rr' \cos\gamma$ and $\cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi - \phi')$. The *multipole expansion* of $1/|x - y|$ is then given by

$$\frac{1}{|x-y|} = \sum_{n=0}^{\infty} \frac{r'^n}{r^{n+1}} P_n(\cos\gamma)\,, \tag{38}$$

where the P_n are the Legendre polynomials. This is not yet a product decomposition since γ is a mix of coordinates of both points. We separate them using the addition theorem for the Legendre polynomials

$$P_n(\cos\gamma) = \sum_{m=-n}^{n} Y_n^{-m}(\theta', \phi') Y_n^m(\theta, \phi)\,,$$

where Y_n^m are the surface harmonics of the first kind. The final product decomposition is given by

$$\frac{1}{|x-y|} = \sum_{n,m} \left(Y_n^{-m}(\theta', \phi') r'^n \right) \left(\frac{Y_n^m(\theta, \phi)}{r^{n+1}} \right). \tag{39}$$

Because of the requirement $r' < r$, we cannot use a single origin for all interactions, and also the interactions of particles close together have to be calculated directly because of the bad convergence of the multipole expansion. This problem is similar to the problem of the short-range forces, where only the interactions with close cells have to be treated. In the case of the multipole method, we treat all cells not adjacent using the multipole expansions, while the adjacent cells are treated using the standard pairwise Coulomb sum. But to gain anything from the multipole expansion, we have to be able to combine the multipole expansion from the remote cells with a multipole expansion of larger clusters. This requires a procedure to shift the origin of a multipole expansion, which can be achieved by simple linear recombinations of the coefficients $\left(Y_n^{-m}(\theta', \phi') r'^n\right)$ and is called a *translation* operation. During the expansion process, it is advantageous to be able to exchange the roles of the particles during a shift, i.e., to translate the origin of the expansion from the vicinity of one of the particles to that of the other which is similarly possible. This procedure is called a *conversion* operation.

Now let us assume that the number of cells per simulation box side is a power of two, $S = 2^L$, where L is termed the number of *levels*. Eight of the cells of the lowest level are always combined to form a division of the simulation box into 2^{L-1} larger cells per side, and so on. We can combine the multipole expansions of eight neighboring level L cells through translations into one multipole expansion for the cell of level $L - 1$ formed by the eight original cells. This procedure can be continued again to obtain multipole expansions for all cells on each level. Once we have arrived at the bottom level $L = 1$ we convert the expansion to local expansions in all three other top-level cells and distribute them up to the higher levels again until for all cells the multipole expansion of the full system is available. The multipole expansions for the top level are used to calculate the electrostatic interaction. A graphical representation of this procedure is shown in Fig. 2.

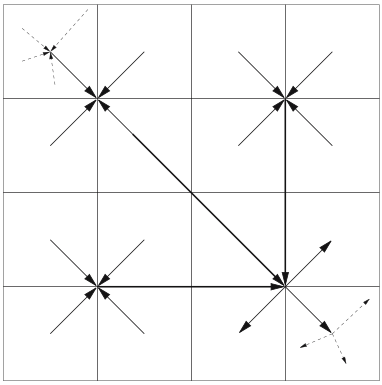

Fig. 2 Graphical representation of the multipole expansion procedure. The *dotted lines* represent the calculation of the multipole expansion and of the energy rsp. forces, and the *solid lines* represent translations and conversionss

Here, only the data flow from one cell to another is shown. In a real simulation this flow occurs for all pairs of cells, and of course the number of cells is much larger.

In pseudo code the energy respective force calculation looks like this:

```
for each particle i
        add contribution of particle p to the
        local multipole expansion of its cell
end
for each l = L...2
        for each cell c of level l
                add multipole expansion to expansion of the
                level l – 1 supercell, translated to its center
end
```

```
for each cell c of level l
        for each cell c' of level l
                add converted contribution of cell c to the
                local expansion for cell c'
for each l = 2...L
        for each cell c of level l
                calculate local multipole expansion from the
                translated expansions of same-level neighbors
                and of the level l – 1 supercell
end
```

The number of terms needed in the multipole expansion only depends on the precession requirement and the number of particles in a cell. At constant density and constant number of particles per cell the number of operations therefore only depends on the loops shown above, which are all either of order N or $S^3(1 + 1/8 + 1/64 + \ldots) = \mathcal{O}(N)$, such that the overall computational order is $\mathcal{O}(N)$. One drawback of the method is that all the intermediate multipole expansions have to be stored, since they are needed in the last loop. This can use a considerable amount of memory, since the number of terms in the multipole expansion can be large.

Tree codes or Barnes–Hut methods [16] work similarly with respect to the generation of the multipole expansions, but use a different algorithm for spreading them to the other cells. This algorithm is not strictly hierarchical and therefore reduces the computational order to $\mathcal{O}(N \log N)$, but considerably less memory is consumed. The algorithm presented so far is only suitable for the nonperiodic case, but modifications for arbitrary periodic boundary conditions exist in terms of modified coefficients of the multipole expansion. These coefficients are obtained from the traditional Ewald sum or the corresponding method for partially periodic boundary conditions. However, due to the increasing number of coefficients, the multipole methods are quite slow, and only for around a million particles do they become comparable in speed to the P^3M method.

2.6 The Lekner Sum

Both the fast multipole methods and the Ewald-type methods adhere to the spherical summation order for the electrostatic interaction. In contrast to this, Lekner came up with a different formula [49], which is based on transforming the force sum along only one coordinate into a rapidly converging sum. Naturally, the summation order is rodwise, along the transformed coordinate first. This alone is no problem, since the difference to the spherical limit is given by a multipole of the dipole moment, very similar to the slabwise summation [50]. Also, the change to forces instead of energy is problematical,

since it is equivalent to changing the order of differentiation and summation in the only conditionally convergent Coulomb sum. Nevertheless, the Lekner summation has been used frequently in recent computer simulations [51–53], mainly because of its extremely good convergence, which makes it very favorable for small numbers of particles.

We need to introduce some change in notation here [15], which will be used frequently in the following calculations. Let the simulation box dimensions be $\lambda_x \times \lambda_y \times \lambda_z$. Then the real-space lattice vectors are

$$n_{klm} := (k\lambda_x, l\lambda_y, m\lambda_z)^T, \quad n_{kl} := n_{kl0}, \quad \text{and} \quad n_k := n_{k0} \tag{40}$$

for $k, l, m \in \mathbb{Z}$. The inverse box dimensions are

$$u_x = \frac{1}{\lambda_x}, \quad u_y = \frac{1}{\lambda_y}, \quad \text{and} \quad u_z = \frac{1}{\lambda_z}\,. \tag{41}$$

The relative particle positions are

$$p_{ij} = (x_{ij}, y_{ij}, z_{ij}) = p_i - p_j\,. \tag{42}$$

For $r = (x, y, z)$ the distances from the origin to the (k, l, m)th periodic image are

$$r_{klm} = |r + n_{klm}| = \sqrt{(x + k\lambda_x)^2 + (y + l\lambda_y)^2 + (z + m\lambda_z)^2} \tag{43}$$
$$r_{kl} = r_{kl0}, \quad r_k = r_{k0}, \quad \text{and} \quad r = r_0$$

and the yz-plane distances are

$$\rho_{lm} = \sqrt{(y + l\lambda_y)^2 + (z + m\lambda_z)^2} \quad \rho_l = \rho_{l0} \quad \text{and} \quad \rho = \rho_0\,. \tag{44}$$

To obtain the Lekner summation formula, one treats the summation along the z-axis as a function of z and expands this function in a Fourier series. This results in the force formula

$$\widetilde{F}_i = 8\pi \sum_{j \neq i} q_i q_j \sum_{l,m \in \mathbb{Z}} \sum_{p \geq 1} p \mathrm{K}_0 \left(2\pi u_x p \rho_{lm}\right) \cos(2\pi u_x p x)\,. \tag{45}$$

The tilde on the force denotes that this force is different from the forces obtained by using the spherical summation limit. The representation is nevertheless well convergent, at least if all charges q_j are well separated in y and z from q_i, i.e., if ρ_{lm} is large enough for all l, m. By spatial symmetry we can choose the coordinate of closest approach as the x-coordinate. For particles close together in all coordinates one has to employ a different method to calculate the electrostatic interaction; see, for example, Sperb [53], who suggests

the following formula

$$\widetilde{F}_i = 8\pi \sum_{j\neq i} q_i q_j \sum_{l,m\in\mathbb{Z}} \left(\frac{z}{\left(\rho_{lm}^2 + z^2\right)^{\frac{3}{2}}} - \sum_{k=0}^{\infty} \binom{-\frac{3}{2}}{k} r^{2k} \left(\psi^{(2k+p)}(1+z) - \psi^{(2k+p)}(1-z) \right) \right), \tag{46}$$

where $\psi^{(m)}$ denotes the mth polygamma function. This version, in contrast, only converges for small ρ_{lm} values, but again is very fast. Since we can choose the Fourier-transformed coordinate freely in Eq. 45, this equation can be used efficiently if at least one of the coordinates is sufficiently large. For particles which are close, we use the alternative form of Eq. 46. Consequently, the formulas are called *far* and *near*, respectively, in the following sections. This concept is also used in the MMM, MMM2D, MMM1D, and ELC methods presented below.

For the Lekner method, upper error bounds can be given easily since the Bessel functions drop essentially exponentially fast, allowing for a simple approximation of the sum by an integral. These error estimates are much less sharp than the error estimates for the Ewald-type methods, but here the error bound only enters logarithmically into the computation time, so that excessive accuracy has only a small impact on the overall performance.

2.7 MMM

Yet another approach to tackle the conditionally convergent Coulomb sum is used for MMM. Instead of defining the summation order, one can also multiply each summand by a continuous factor $c(\beta, r_{ij}, n_{klm})$ such that the sum is absolutely convergent for $\beta > 0$, but $c(0, ., .) = 1$. The energy is then defined as the limit $\beta \to 0$ of the sum, i.e., β is an artificial convergence parameter. For a convergence factor of $\mathrm{e}^{-\beta n_{klm}^2}$ the limit is the same as the spherical limit, and one can derive the classical Ewald method quite conveniently through this approach [50]. To derive the formulas for MMM, one has to use a different convergence factor, namely $\mathrm{e}^{-\beta|r_{ij}+n_{klm}|}$, which defines the alternative energy

$$\widetilde{E} = \frac{1}{2} \lim_{\beta\to 0} \sum_{k,l,m\in\mathbb{Z}} {\sum_{i,j=1}^{N}}' \frac{q_i q_j \,\mathrm{e}^{-\beta|p_{ij}+n_{klm}|}}{|p_{ij}+n_{klm}|}. \tag{47}$$

The limit $\widetilde{E}$ exists, but differs by some multiple of the square of the dipole moment from the spherical limit as obtained by the Ewald summation [50]. From the physical point of view, the Coulomb interaction is replaced by a screened Coulomb interaction with screening length $1/\beta$. $\widetilde{E}$ is then the energy in the limit of infinite screening length. But because of the conditional convergence

of the electrostatic sum, this is not necessarily the same as the energy of an unscreened system. Since the difference to the Ewald methods only depends on the dipole moment of the system, the correction can be calculated easily in linear time, and can be ignored with respect to accuracy and computational time.

Starting from this convergence factor approach along the lines of the Lekner sum, R. Strebel and R. Sperb constructed a method of computational order $\mathcal{O}(N \log N)$, MMM [54]. The favorable scaling is obtained, very much as in the Ewald case, by technical tricks in the calculation of the far formula. The far formula has a product decomposition and can be evaluated hierarchically, similar in spirit to the fast multipole methods.

The goal is to find a pairwise interaction ϕ satisfying

$$\widetilde{E} = \frac{1}{2} \sum_{i,j=1}^{N} q_i q_j \phi(x_{ij}, y_{ij}, z_{ij}) \,. \tag{48}$$

For particles sufficiently separated in the z-axis, one can Fourier transform the outer sum in Eq. 47 along both k and l. The resulting formula is singular in $\beta \to 0$, however, the singularity is independent of the particle coordinates, and therefore vanishes due to charge neutrality and the summation over all charges. For the remaining part of the Fourier-transformed equation, the limit $\beta \to 0$ can be taken straight forward. One obtains a representation in the form of Eq. 48, with the far formula for ϕ given by

$$\begin{aligned} \phi(x,y,z) = u_x u_y \sum_{p,q\neq 0} \frac{\mathrm{e}^{2\pi f_{pq} z} + \mathrm{e}^{2\pi f_{pq}(\lambda_z - z)}}{f_{pq}\left(\mathrm{e}^{2\pi f_{pq}\lambda_z} - 1\right)} \mathrm{e}^{2\pi \mathrm{i} u_y q y} \mathrm{e}^{2\pi \mathrm{i} u_x p x} \\ + 2\pi u_x u_y \left(u_z z^2 - z + \frac{\lambda_z}{6}\right) , \end{aligned} \tag{49}$$

where

$$\begin{aligned} f_{pq} = \sqrt{(u_x p)^2 + (u_y q)^2}, \quad f_p = u_x p, \quad f_q = u_x q, \\ \omega_p = 2\pi u_x p, \quad \text{and} \quad \omega_q = 2\pi u_y q \,. \end{aligned} \tag{50}$$

The advantage of this formula is that it allows for product decomposition into components of the particles. For example,

$$\mathrm{e}^{2\pi f_{pq} z} = \mathrm{e}^{2\pi f_{pq}(z_i - z_j)} = \mathrm{e}^{2\pi f_{pq} z_i} \mathrm{e}^{-2\pi f_{pq} z_j}$$

etc. Therefore, one just has to calculate the sum over all these exponentials on the left side and on the right side, and multiply them together, which can be done in $\mathcal{O}(N)$ computation time. As can be easily seen, the convergence of the series is excellent as long as z is sufficiently large. By symmetry, one can choose as z the coordinate with the largest distance in order to optimize the convergence. Similar to the Lekner sum, we need a different formula if all co-

ordinates are small, i.e., for particles close to each other. For sufficiently small $u_y\rho$ and $u_x x$ we obtain the near formula as

$$\begin{aligned}\tilde{\phi}(x,y,z) =& 2u_x u_y \sum_{p,q>0} \frac{\cosh(2\pi f_{pq} z)}{f_{pq}\left(\mathrm{e}^{2\pi f_{pq}\lambda_z} - 1\right)} \mathrm{e}^{2\pi \mathrm{i} u_y q y} \mathrm{e}^{2\pi \mathrm{i} u_x p x} \\ &+ 4u_x \sum_{l,p>0} \left(\mathrm{K}_0(2\pi u_x p \rho_l) + \mathrm{K}_0(2\pi u_x p \rho_{-l})\right) \cos(2\pi u_x p x) \\ &- 2u_x \sum_{n\geq 1} \frac{b_{2n}}{2n(2n)!} \mathrm{Re}\left((2\pi u_y (z + \mathrm{i}y))^{2n}\right) \\ &+ u_x \sum_{n\geq 0} \binom{-\frac{1}{2}}{n} \frac{\left(\psi^{(2n)}(1 + u_x x) + \psi^{(2n)}(1 - u_x x)\right)}{(2n)!} \rho^{2n} \\ &- 2\log(4\pi)\,.\end{aligned} \tag{51}$$

Note that this time we calculate $\tilde{\phi}$ instead of ϕ, which denotes that we omit the contribution of the primary simulation box. This is very convenient, as it includes the self-energy case, and makes $\tilde{\phi}$ a smooth function. To obtain ϕ one has to add the $1/r$ contribution of the primary box, except for the self-interaction of a particle. The self-energy term is given by

$$\begin{aligned}\tilde{\phi}(0,0,0) =& 2u_x u_y \sum_{p,q>0} \frac{1}{f_{pq}\left(\mathrm{e}^{2\pi f_{pq}\lambda_z} - 1\right)} + 8u_x \sum_{l,p>0} \mathrm{K}_0(2\pi u_x \lambda_y p l) \\ &+ 2u_x \psi^{(0)}(1) - 2\log(4\pi)\,.\end{aligned} \tag{52}$$

Equation 51 is derived using the same convergence factor approach as used for Eq. 49, and consequently the same singularity in β is obtained (and omitted). This is important, since otherwise the charge neutrality argument does not hold and the limit $\beta \to 0$ cannot be performed.

By a simple implementation the simulation box is segmented in $B = S^3$ smaller boxes or cells. For all particles the interactions within the cell itself and the 26 neighboring cells are treated using the near formula, while for the rest the far formula is used (see Fig. 3).

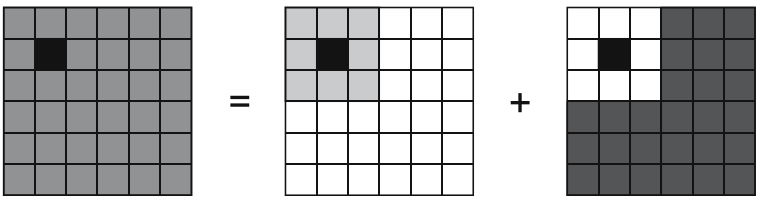

Fig. 3 Two-dimensional representation of the splitting of the interactions calculation using near and far formulas

The interactions of the black cell with the light gray cells is done via the near formula, while all the dark gray cells are treated using the far formula. One first determines the product decomposition components for each cell and then adds them up over all pairs of cells which are not neighbors.

Of course, the infinite sums have to be cut off at some radius. The cutoff radii can be determined using error formulas, for example the maximal error, denoted by τ_ϕ, of ϕ. It is calculated by summing the far formula up to a pq cutoff of R, hence

$$\tau_\phi \approx \frac{e^{-2\pi Rr}}{r}, \tag{53}$$

where r is the distance between the particles. The error formula reflects the condition that two particles have to be sufficiently separated to efficiently use the far formula. The error estimates are similar to those that will be presented for MMM2D later.

Using the algorithm described above, the minimal distance of two particles calculated with the far formula is λ_z/S. Therefore, for a constant pairwise error the Fourier-space cutoff R has to be chosen proportional to S. This leads to a calculation time for the far formula of $\mathcal{O}(NS^2)$. The near formula has to be used for $\mathcal{O}(N^2S^{-3})$ particle pairs. Since the calculation time for the near formula is practically parameter independent, this is also the scaling of the calculation time. The total computation time has a minimum for $S \sim N^{1/5}$, resulting in an overall computation time scaling of $\mathcal{O}(N^{7/5})$.

To decrease the computational effort of MMM down to $\mathcal{O}(N \log N)$ we use the periodicity of the axis which dominates the error, i.e., in which the particles are closest. Since the far-formula error of MMM2D is dominated by the error in the nonperiodic direction, the following cannot be transferred to MMM2D, and is unique to the MMM method.

We assume that the number of cells per side is a power of 2, i.e., $S = 2^L$. The main idea is to increase λ_z/S not by decreasing S, but rather by increasing λ_z. This is possible due to the periodic boundary conditions, as we will see now.

Again, the idea will be presented graphically in two dimensions, and again we calculate the interactions with the small black cell. First, the primary simulation cell will be divided into small cells along the x-coordinate (see Fig. 4).

The particles in the right part are far away and can be calculated using the far formula with a reasonably small cutoff.

The other half will now be calculated using a cell length of $\lambda_x/2$. This introduces artificial particles in the right part, which are just copies of the particles in the left side. Their contribution can be subtracted easily together with the contribution of the real particles in the right half (see Fig. 5).

Now we are left with a cell containing roughly $N/2$ particles (provided the simulation cell is filled homogeneously) and cell dimensions $\lambda_x/2 \times \lambda_y \times \lambda_z$. For this system we apply the same trick again to one of the other axes, e.g.,

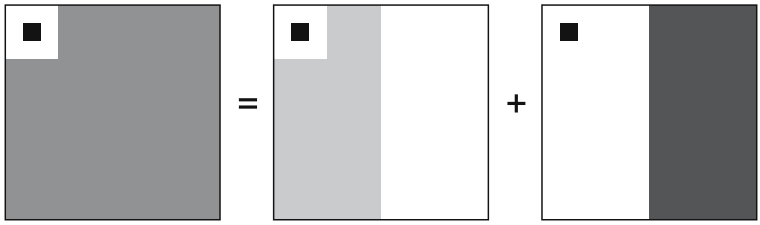

Fig. 4 Division of the primary simulation cell into nearby and far away regions

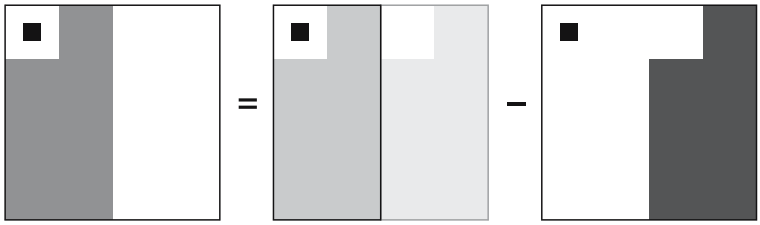

Fig. 5 Calculation for particles in the left part in Fig. 4

y, then to z and again to x, and so on, until the calculation using the near formula is more efficient than another subdivision. The subdivision occurs in the coordinate which occurs in the exponential. Therefore, the shift of the particle coordinates is actually only a multiplication.

This algorithm can be further optimized with respect to computation time, but even in the present form one can see that its implementation is more demanding than the Ewald or FMM methods. The implementation is even more complex, since the presented code should work with all three coordinates symmetrically for reuse in the subdivision steps. The calculation time for each of the subdivision steps is proportional to half the number of particles left, i.e., $N/2, N/4, N/8, \ldots$. To maintain a constant calculation time of the near formula, one has to ensure that $2^L \sim N$ or $L \sim \log N$. Therefore, the overall computation time is $\mathcal{O}(NL) = \mathcal{O}(N \log N)$.

Upper error bounds can be found easily by approximating the sums with integrals (see [54]). As for the Lekner sum, the additional accuracy has to be paid for with only a small decrease of computational performance; therefore, MMM is the method of choice if high accuracy is required.

3 Two Periodic Dimensions

Thin polyelectrolyte films or interactions of charged species with membranes require a different periodicity, where only two dimensions are pe-

riodically replicated, while the third one has finite extent h ($2D + h$ geometry). For this geometry the original Ewald sum has to be modified. For such systems Ewald-based formulas are only slowly convergent, have mostly $\mathcal{O}(N^2)$ scalings, and no "a priori" error estimates exist [55]. A fast Ewald-based method has been put forward in Ref. [56], where the Fourier sum is specially treated with an integral approximation. However, in our opinion, this introduces noncontrollable errors, which makes the method useless [57]. The Lekner method, which is based on a resummation of the force sum, can also be used; however, it has an $\mathcal{O}(N^2)$ scaling, and special care has to be taken for particles which are nearby in the nonperiodic dimension (see also the discussion of Mazars [51] on accuracy problems). As discussed before, one can use multipole methods; however, this approach is only favorable for huge numbers of particles and not suited to typical simulations.

Recently we proposed a new method called MMM2D [24, 58], which has an $\mathcal{O}(N^{5/3})$ complexity and full error control that is based on a convergence factor approach similar to MMM [12]. In two dimensions the convergence factor based methods and the Ewald sum methods yield exactly the same results. However, this will still only allow simulations including up to a few thousand charges due to the power law scaling.

Because of the superior scaling of the methods for fully periodic systems, there have been early attempts to use the full 3D Coulomb sum also for slab problems. The main idea is to fill only parts of the simulation box with charges and to leave some space empty, in an attempt to decouple the interactions in the third dimension [59–61]. Since each image layer is globally neutral, one hopes that their interactions decay as they become more and more distant, i.e., as the size of the gap is increased. In this way one could make use of any advanced 3D Ewald implementation (see also Ref. [62] for a variant of this idea). In the following discussion we present the ELC approach which allows one to subtract the contributions of the image layers again, thereby allowing for much smaller gap sizes and accelerating the method considerably.

3.1 The MMM2D Method

Since the ELC method is based on the MMM2D method, and because MMM2D is computationally faster than any 3D method in combination with ELC for small numbers of particles, we start with a short description of MMM2D. In analogy to MMM, we use a representation with an appropriate pairwise interaction ϕ, given by Eq. 48. Unlike the fully periodic case, in partially periodic boundary conditions the convergence factor energy $\widetilde{E}$ is equal to the energy obtained from the standard spherical summation. The pairwise

interaction ϕ is given by the MMM2D far formula

$$\begin{aligned}\phi(x,y,z) = 4u_x u_y \sum_{p,q>0} \frac{\mathrm{e}^{-2\pi f_{pq}|z|}}{f_{pq}} \cos(\omega_p x)\cos(\omega_q y) \\ + 2u_x u_y \left(\sum_{q>0} \frac{\mathrm{e}^{-2\pi f_q|z|}}{f_q} \cos(\omega_q y) + \sum_{p>0} \frac{\mathrm{e}^{-2\pi f_p|z|}}{f_p} \cos(\omega_p x) \right) \\ - 2\pi u_x u_y |z| \,,\end{aligned} \tag{54}$$

using the same notations as for Eq. 49. Note that the far formula is well convergent for $|z| \gg 0$. The near formula is

$$\begin{aligned}\tilde{\phi}(x,y,z) = 4u_x \sum_{l,p>0} \left(\mathrm{K}_0(\omega_p \rho_l) + \mathrm{K}_0(\omega_p \rho_{-l})\right) \cos(\omega_p x) \\ - 2u_x \sum_{n\geq 1} \frac{b_{2n}}{2n(2n)!} \mathrm{Re}\left((2\pi u_y (z + \mathrm{i}y))^{2n}\right) + \sum_{k=1}^{N_\psi - 1} \left(\frac{1}{r_k} + \frac{1}{r_{-k}}\right) \\ - u_x \sum_{n\geq 0} \binom{-\frac{1}{2}}{n} \frac{\left(\psi^{(2n)}(N_\psi + u_x x) + \psi^{(2n)}(N_\psi - u_x x)\right)}{(2n)!} (u_x \rho)^{2n} \\ - 2u_x \log\left(4\pi \frac{u_y}{u_x}\right) \,.\end{aligned} \tag{55}$$

The algorithmic tricks used in MMM cannot be applied in two-dimensionally periodic boundary conditions, since these tricks require the periodicity of all coordinates. However, the far formula still allows for a product decomposition and can be evaluated in linear time with respect to the number of particles with constant cutoffs.

To draw advantage from this linear computation time scaling, one splits the system into B equally sized layers along the nonperiodic z-coordinate. Now for all particles in one slice, the interactions with the neighboring slices are calculated using the near formula, while for the further distant slices the far formula is used. This ensures that the far formula is used only for particle pairs more distant than h/B, where h is the total system height. The far formula is evaluated in a similar fashion, as will be presented below for the ELC method, although some more involved bookkeeping is necessary. This is due to the fact that the far contribution is only needed for particles from certain slices, whereas the ELC term has to be calculated for all particles.

Assuming that N/B particles are located in every slice, one can show that the far formula scales like $\mathcal{O}(B^2 N) + \mathcal{O}(B^3)$. The time for the calculation of the near formula for a single particle pair is nearly constant, leading to a scaling of $\mathcal{O}(N(2N/B))$. Minimizing the total time with respect to the number of slices B leads to $B \propto N^{1/3}$, yielding an asymptotic optimal computational time of $\mathcal{O}(N^{5/3})$. Although we know the optimal dependence of B on N, this pa-

rameter still has to be tuned to the underlying hardware. This can, however, be done by a simple trial and error algorithm at runtime.

MMM2D allows for error bounds similar to MMM or the Lekner method, and it also inherits the advantageous scaling of the computation time with the error bound. Due to the splitting of the interaction calculation into to different formulas, the error distribution of MMM2D is not uniform (see Fig. 6). Therefore, averaging error estimates such as the rms force error (Eq. 26) can be misleading for this type of method. Therefore, we present the maximal pairwise error, i.e., an error estimate for the maximal error in the MMM2D pair potential. From this maximal pairwise error, one obtains an upper bound to the rms force error by $\Delta F \leq \sqrt{3/N} \sum_{i=1}^{N} q_i^2 \tau_{\mathrm{F}}$, where τ_{F} is the maximal pairwise force error. However, the real rms force error is normally an order of magnitude lower. The derivation of the error estimates is straightforward, and we present only the results here. For the maximal pairwise force error of the far formula, one obtains:

$$|\Delta F| \leq \tau_{\mathrm{F}}^{\mathrm{far}} = \frac{\mathrm{e}^{-2\pi R|z|}}{|z|} \left(2\pi R + 2(u_x + u_y) + \frac{1}{|z|} \right), \tag{56}$$

where R is the cutoff radius in $(u_x p, u_y q)$ of the far formula. A similar expression can also be found for the energy error.

For the near formula, one has separate error estimates for the sum of Bessel terms, the complex sum, and the polygamma sum. For the Bessel sum,

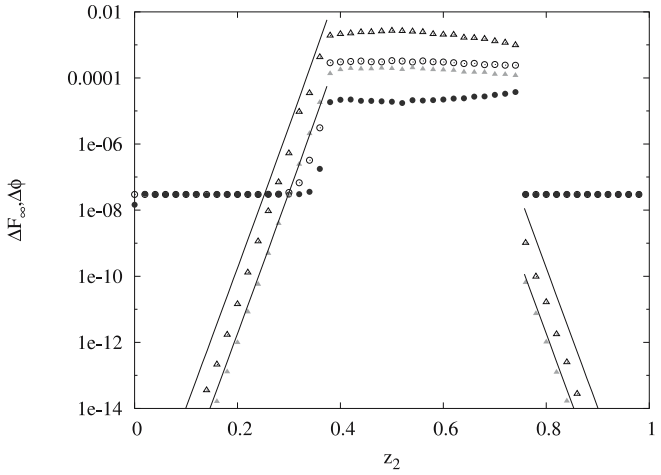

Fig. 6 Absolute force component error $\Delta F_\infty = \max(|\Delta F_x|, |\Delta F_y|, |\Delta F_z|)$ (*triangles*) and energy error $\Delta\phi$ (*circles*) as a function of the z position of a randomly placed particle in a $1 \times 1 \times 1$ box; a second particle was fixed at $z = 0.5$. For both the force and energy error the *open symbols* give the maximal error that occurred, and the *filled symbols* the average error. The *solid line* shows the force vs energy error estimates for MMM2D

the maximal force error is

$$\tau_{\mathrm{F}}^{\mathrm{Bessel}} = 16\pi u_x^2 \mathrm{K}_1(\lambda_y L)\left(\frac{\mathrm{e}^{\pi u_x \lambda_y}}{\pi u_x \lambda_y}\left(\frac{L+u_y}{\pi u_x} - 1\right) + \sum_{p=1}^{\lceil \frac{L}{\pi u_x} \rceil - 1} p\,\mathrm{e}^{-\pi u_x \lambda_y p}\right) \tag{57}$$

where the summation takes place over all (p, l), where $0 < p < \frac{L}{\pi u_x}$ and $0 < l < \frac{L}{\omega_p} + 1$. For the summation of the complex sum up to order N, one obtains the maximal error bound

$$\tau_{\mathrm{F}}^{\mathrm{Bernoulli}} = 16 u_x u_y (u_y \rho)^{2N-1} \leq 16\sqrt{2} u_x u_y 2^{-N}\,. \tag{58}$$

The polygamma sum finally is a sum of alternating coefficients with monotonously decreasing value, and therefore by Leibniz' criterion, the absolute value of the last term taken into account is an upper error bound.

If the error bound for both the near and far formulas is not chosen low enough, e.g., two orders of magnitude below the thermal fluctuations, the errors may lead to a heating of the system within particle clusters, and to a cooling near the system borders. This is in general no problem, since precision is inexpensive in the MMM family of methods, so that the computation time does not suffer much from the increased precision requirements, and one has a fast and reliable method for the calculation of the electrostatic interaction in two-dimensionally periodic systems.

3.2 The ELC Method

The MMM2D method is well-suited for computations with up to a few hundred particles. For thousands or even hundreds of thousands of particles, however, this method cannot be used. Until recently, the only viable way out was to use a 3D method, with a large empty gap between the artificial image layers.

Since no error bounds were known for this approach, it had to be checked on a trial and error basis. We greatly improved this situation by deriving a term, called electrostatic layer correction (ELC), which exactly calculates the interactions due to the unwanted layers [25, 26]. These can then be subsequently subtracted, and in this way the necessary gap length can be largely reduced. The combination of that term with any three-dimensional summation method with slabwise summation order will yield the exact electrostatic energy. Since the change in the summation order is done by adding a very simple term, any three-dimensional summation method with the standard spherical summation order can be used. The ELC term can be easily evaluated linearly in time with the number of charges, hence the whole method scales like the underlying standard summation method. We also developed an error formula for the maximal pairwise error in the energy and forces of

the layer correction term; hence, the precision of this method can be tuned to any desired value when used in conjunction with other error estimates for the standard summation method [22, 35]. Using, for example, the P^3M method one obtains an $N \log N$ scaling with well-controlled errors also for the $2D + h$ geometry, which up to now seems to be the optimal choice.

Let us start by describing how to change the summation order. Consider this time a system of N particles with charges q_i and positions $p_i = (x_i, y_i, z_i)$ that reside in a box of edges $\lambda_x \times \lambda_y \times h$, where $h = \max_{i,j}|z_i - z_j|$ is the maximal z-distance of two particles. The basic idea is to expand this slab system in the nonperiodic z-coordinate to a system with periodicity in all three dimensions. More precisely, the original box of size $\lambda_x \times \lambda_y \times h$ is placed inside a box of size $\lambda_x \times \lambda_y \times \lambda_z$ where $\lambda_z \gg h$ is sufficiently large. Then this box is replicated periodically in all three dimensions. The result is a three-dimensional periodic system with empty space regions ("gaps") of height $\delta := \lambda_z - h$ (see Fig. 7); δ is called the gap size.

Since the electrostatic potential is only finite if the total system is charge neutral, the additional image layers (those layers above or below the original slab system) are charge neutral, too. Now let us consider the nth image layer which has an offset of $n\lambda_z$ to the original layer. If $n\lambda_z$ is large enough, each particle of charge q_j at position $(x_j, y_j, z_j + n\lambda_z)$ and its replicas in the x, y-plane can be viewed as constituting a homogeneous charged sheet of charge density $\sigma_j = \frac{q_j}{\lambda_x \lambda_y}$. The potential of such a charged sheet at distance z is $2\pi\sigma_j|z|$. Now we consider the contribution from a pair of image layers located at $\pm n\lambda_z$, $n > 0$ to the energy of a charge q_i at position (x_i, y_i, z_i) in the central layer. Since $|z_j - z_i| < n\lambda_z$, we have $|z_j - z_i + n\lambda_z| = n\lambda_z + z_j - z_i$ and $|z_j - z_i - n\lambda_z| = n\lambda_z - z_j + z_i$, and hence the interaction energy from those two image layers with the charge q_i vanishes by charge neutrality:

$$2\pi q_i \sum_{j=1}^{N} \sigma_j(|z_j - z_i + n\lambda_z| + |z_j - z_i - n\lambda_z|) = 4\pi q_i n\lambda_z \sum_{j=1}^{N} \sigma_j = 0\,. \tag{59}$$

The only errors occurring are those coming from the approximation of assuming homogeneously charged, infinite sheets instead of discrete charges.

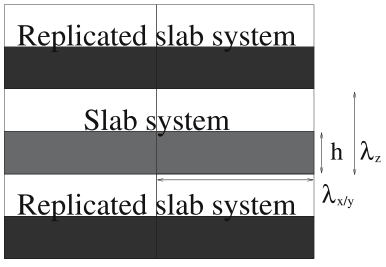

Fig. 7 Schematic representation of a fully periodically replicated slab system

This assumption should become better when increasing the distance $n\lambda_z$ from the central layer.

However, in a naive implementation, even large gap sizes will result in large errors [61]. This is due to the order of summation for the three-dimensional Coulomb sum, which is spherical by convention. This order implies that with increasing shell cutoff S the number of image shells grows faster than the number of shells of the primary layer, namely $\mathcal{O}(S^3)$ versus $\mathcal{O}(S^2)$ (see Fig. 8a). In other words, we include the unwanted terms faster than the actually wanted terms. Also, the image layers are not really infinite charged sheets but are truncated due to the cutoff. Yeh and Berkowitz [61] showed that this problem can be solved by changing to a slabwise summation order, using the old results of Smith [50]. As already discussed in Sect. 2.1, Smith has shown that by adding to the Coulomb energy the term

$$E_c = 2\pi M_z^2 - \frac{2\pi M^2}{3}, \tag{60}$$

where $M = \sum q_i p_i$ is the total dipole moment, one obtains the result of a slabwise summation instead of the spherical limit [50]. Slabwise summation refers to the sum $\sum_{|n|\geq 0} E_l(n)$, where $E_l(n)$ denotes the energy, calculated in circular summation order, resulting from the image layer with shift $n\lambda_z$ in the z-coordinate. Technically this is the order where we first treat the original layer and then add the image layers grouped in symmetrical pairs (see Fig. 8b). Obviously this summation order fits much better to the charged sheet argument given above. Although this is a major change in the summation order, the difference given by Eq. 60 is a very simple term. In fact, Smith showed that changes of the summation order always result in a difference that depends only on the total dipole moment.

Applying this slabwise summation order, Yeh and Berkowitz showed that using a gap size of some multiples of the box size is normally sufficient to

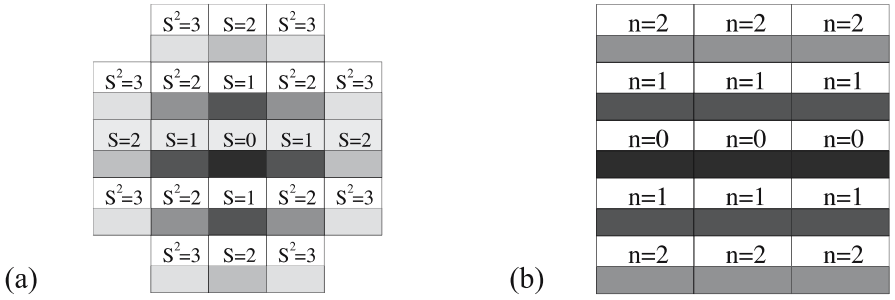

Fig. 8 Schematic views of different summation orders. **a** Schematic view of the spherical summation order. S is the length of the box offset. **b** Schematic view of the slabwise summation order. n is the z offset of the box, the spherical summation order in the x,y-plane is not shown

obtain a moderately accurate result. Therefore, the result of a standard three-dimensional summation method plus the shape-dependent term given by Eq. 60, which we refer to as a *slabwise method*, can be used to obtain a good approximation to the result for the slab geometry with the same computational effort as for the underlying three-dimensional summation method (no matter whether a simple or sophisticated method is used). One drawback is that no theoretical estimates exist for the error introduced by the image layers. Therefore, one might be forced to use even larger gaps to assure that no artifacts are produced by the image layers. One deducible artifact is that the pairwise error will be position dependent. Particles in the middle of the slab will see no effect of the image layers due to symmetry, and particles near the surface will encounter, for the same reason, the largest errors, which is definitely an unwanted feature for studying surface effects. Therefore, averaging error measures like the commonly used rms force error should not be applied without additional checks for the particles near the surfaces.

The other drawback is that normally the box now will have a significantly larger λ_z/λ_x versus λ_z/λ_y. But at least for Ewald-type methods the computation time is proportional to this fraction. This is easy to see as the number of k-space vectors in the z direction must be proportional to λ_z to maintain a fixed resolution and therefore error. It is verified experimentally that a gap of several h is needed, therefore the computation time more than doubles.

We will now show how the ELC term can be calculated with a computation time proportional to the number of particles. Since this is as fast as any method for 3D periodic boundary conditions, the ELC term will not dominate the computation time scaling for any 3D method. Moreover, it has a small prefactor, so that a 3D method combined with ELC can be significantly faster. For more details, consult the original articles [25, 26].

The method relies on the far formula of MMM2D (Eq. 54). For the following analysis there is no special restriction on h except for $h < \lambda_z$, which is true even if the $\lambda_x \times \lambda_y \times \lambda_z$ box is completely filled.

We start with a formal definition of the Coulomb energy of the slab system

$$E = \frac{1}{2} \sum_{S=0}^{\infty} \sum_{k^2+l^2=S} \sum_{i,j=1}^{N}{}' \frac{q_i q_j}{|p_i - p_j + n_{kl}|} \tag{61}$$

where n_{kl} denotes the lattice vector (see notations for Eq. 40). As usual, the prime on the inner summation indicates the omission of the self-interaction $i = j$ in the primary box $k = l = 0$ (i.e., the singular case). For the surrounding dielectric medium we assume vacuum boundary conditions.

We now expand the system to a fully three-dimensional periodic system, where λ_z determines the period in the z-coordinate, as in the previous section. We can rewrite the energy as

$$E = E_s + E_c + E_{lc}\,, \tag{62}$$

where

$$E_s = \frac{1}{2}\sum_{S=0}^{\infty} \sum_{\substack{k,l,m\in\mathbb{Z}\\ k^2+l^2+m^2=S}} \sum_{i,j=1}^{N}{}' \frac{q_i q_j}{|p_i - p_j + n_{klm}|} \tag{63}$$

denotes the standard three-dimensional Coulomb sum with spherical limit. To evaluate this expression one can use any of the efficient algorithms, starting with the classical Ewald summation up to modern methods like fast multipole methods [17] or mesh-based algorithms [10]. E_c again denotes the shape-dependent term given by Eq. 60, and finally

$$E_{lc} = -\frac{1}{2}\sum_{\substack{T\in\mathbb{Z}\\ T>0}} \sum_{m=\pm T} \sum_{S=0}^{\infty} \sum_{\substack{k,l\in\mathbb{Z}\\ k^2+l^2=S}} \sum_{i,j=1}^{N} \frac{q_i q_j}{|p_i - p_j + n_{klm}|}\,. \tag{64}$$

denotes the contribution of the image layers, for which we are going to derive a new expression below.

We start with the expression for the energy induced by an image layer at z-offset $m \neq 0$:

$$E_{lc}(m) = -\frac{1}{2}\sum_{S=0}^{\infty} \sum_{k^2+l^2=S} \sum_{i,j=1}^{N} \frac{q_i q_j}{|p_i - p_j + n_{klm}|}\,. \tag{65}$$

It can be shown rigorously, although this is nontrivial, that

$$E_{lc}(m) = -\frac{1}{2}\sum_{S=0}^{\infty} \sum_{k^2+l^2=S} \sum_{i,j=1}^{N} \frac{q_i q_j \,\mathrm{e}^{-\beta|p_i-p_j+n_{klm}|}}{|p_i - p_j + n_{klm}|}\,, \tag{66}$$

using the same convergence factor approach as for MMM2D, which is exact for two-dimensionally periodic systems.

In Refs. [24, 63] one can find a proof for this equation and an efficient way of calculating E_{lc} for charge neutral systems. We do not want to go through the full derivation again; it consists of applying Poisson's summation formula along both periodic coordinates and performing the limit $\beta \to 0$ analytically. One obtains

$$E_{lc}(m) = -\frac{1}{2}\sum_{i,j=1}^{N} q_i q_j \phi(p_i - p_j + n_{00m})\,, \tag{67}$$

where ϕ is the MMM2D far formula given by Eq. 54, and is an artificial pairwise potential that yields the total Coulomb energy, and its derivative produces the pairwise forces for the periodic system.

For now we only have a formula for the contribution of one image layer, so we still have to sum over all m. This task can be performed analytically, and combining the terms for $\pm m$, we obtain

$$E_{lc}(m) = \sum_{i,j=1}^{N} q_i q_j \psi(p_i - p_j)\,, \tag{68}$$

where

$$\begin{aligned}\psi(x,y,z) = 4u_x u_y \sum_{p,q>0} \frac{\cosh(2\pi f_{pq} z)}{f_{pq}(\mathrm{e}^{2\pi f_{pq}\lambda_z} - 1)} \cos(\omega_p x)\cos(\omega_q y) \\ + 2u_x u_y \sum_{p>0} \frac{\cosh(2\pi f_p z)}{f_p(\mathrm{e}^{2\pi f_p \lambda_z} - 1)} \cos(\omega_p x) \\ + 2u_x u_y \sum_{q>0} \frac{\cosh(2\pi f_q z)}{f_q(\mathrm{e}^{2\pi f_q \lambda_z} - 1)} \cos(\omega_q y)\,.\end{aligned} \tag{69}$$

The forces can be obtained from this analysis by simple differentiation, since the sums are absolutely convergent. Although the form in Eq. 69 has a much better convergence than the original form in Eq. 64, its main advantage is a linear computation time with respect to the number of particles N. To see this, the equation has to be rewritten using the addition theorems for the cosine and the hyperbolic cosine. First, one calculates the eight terms

$$\chi_i^{(c/s,c/s,c/s)} = q_i \cosh / \sinh(2\pi f_{pq} z_i) \cos / \sin(\omega_p x_i) \cos / \sin(\omega_q y_i) \tag{70}$$

for the (p,q) vectors with $p, q \neq 0$ and then the eight terms

$$\begin{aligned}\chi_i^{(x,c/s,c/s)} &= q_i \cosh / \sinh(2\pi f_p z_i) \cos / \sin(\omega_p x_i), \\ \chi_i^{(y,c/s,c/s)} &= q_i \cosh / \sinh(2\pi f_q z_i) \cos / \sin(\omega_q y_i)\end{aligned} \tag{71}$$

for the $(p,0)$ and $(0,q)$ vectors, where the indices in the obvious way determine which of the functions cosine (hyperbolicus) or sine (hyperbolicus) are used. Then we evaluate

$$\begin{aligned}E_{lc} =& 4u_x u_y \sum_{p,q>0} \sum_{i,j=1}^{N} \frac{1}{(\mathrm{e}^{2\pi f_{pq}\lambda_z} - 1) f_{pq}} \\ &\times \left(\chi_i^{(ccc)}\chi_j^{(ccc)} + \chi_i^{(csc)}\chi_j^{(csc)} + \chi_i^{(ccs)}\chi_j^{(ccs)} + \chi_i^{(css)}\chi_j^{(css)} \right. \\ &\left. - \chi_i^{(scc)}\chi_j^{(scc)} - \chi_i^{(ssc)}\chi_j^{(ssc)} - \chi_i^{(scs)}\chi_j^{(scs)} - \chi_i^{(sss)}\chi_j^{(sss)} \right) \\ &+ 2u_x \sum_{p>0} \sum_{i,j=1}^{N} \frac{1}{(\mathrm{e}^{2\pi f_p \lambda_z} - 1) f_p} \\ &\times \left(\chi_i^{(xcc)}\chi_j^{(xcc)} + \chi_i^{(xcs)}\chi_j^{(xcs)} - \chi_i^{(xsc)}\chi_j^{(xsc)} - \chi_i^{(xss)}\chi_j^{(xss)} \right) \\ &+ 2u_y \sum_{q>0} \sum_{i,j=1}^{N} \frac{1}{(\mathrm{e}^{2\pi f_q \lambda_z} - 1) f_q} \\ &\times \left(\chi_i^{(ycc)}\chi_j^{(ycc)} + \chi_i^{(ycs)}\chi_j^{(ycs)} - \chi_i^{(ysc)}\chi_j^{(ysc)} - \chi_i^{(yss)}\chi_j^{(yss)} \right).\end{aligned} \tag{72}$$

Similar expansions using the same 16 terms can also be found for the forces. Obviously this has linear computation time with respect to the number of particles, as the only summations over all the particles occur in the χ_*. The infinite summation over (p, q) can be cut off at some value to achieve the desired accuracy. For a cutoff radius of R in the $(u_x p, u_y q)$ space, an upper bound to the maximal pairwise force error is given by

$$\tau_{\mathrm{F}} := \frac{1}{2(\mathrm{e}^{2\pi R\lambda_z} - 1)} \left(\left(2\pi R + 2(u_x + u_y) + \frac{1}{\lambda_z - h} \right) \frac{\exp(2\pi Rh)}{(\lambda_z - h)} + \left(2\pi R + 2(u_x + u_y) + \frac{1}{\lambda_z + h} \right) \frac{\exp(-2\pi Rh)}{(\lambda_z + h)} \right). \tag{73}$$

For a discussion of the maximal pairwise error estimate, see the section on MMM2D. As for MMM2D, a similar expression can be found for the maximal pairwise energy error (for details, see Arnold et al [25]). Using a zero cutoff corresponds to the Yeh and Berkowitz method, for which the ELC error formulas are valid as well. From these we found that the error for these methods decays exponentially in λ_z. However, the errors are not uniformly distributed over the slab; they are worst at the surfaces of the slabs, hence the maximal pairwise error should be used instead of the usual rms errors of Eq. 26 (see Fig. 9).

To summarize, the main profits of the ELC term are that it scales as the number N of particles and has a rigorous error bound. Moreover, this error bound can be used to estimate the size of the image layer contribution and therefore gives a bound on the error introduced by slabwise methods, as proposed by Yeh and Berkowitz [61]. In the second paper on ELC [26] we considered in detail the implementation of the layer correction for the standard Ewald and the $\mathrm{P}^3\mathrm{M}$ methods. There we also derived anisotropic Ewald error formulas, gave some fundamental guidelines for optimization, demonstrated the accuracy, and gave error formulas and computation times for typical systems.

4
The Coulomb Sum in One Periodic Dimension

Boundary conditions where only one dimension is periodic and the other two are open or finite appear in physical situations such as electronic structures of supported structures on metal surfaces, e.g., steps and atomic chains in one dimension, or in systems which can model a charged stiff polymer or DNA piece where, for avoiding end effects, the rod is made infinitely long [64–66]. It was only very recently that a 1D Ewald method (EW1D) for these systems was developed [67, 68], although the Lekner method [11] de-

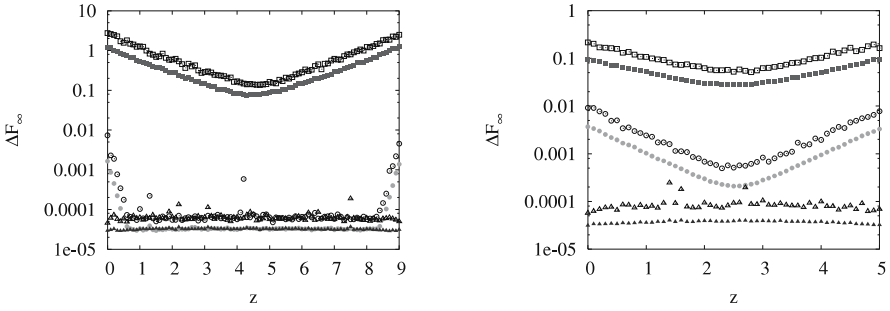

Fig. 9 Absolute force error ΔF_∞ of the ELC term as a function of the particle z-coordinate for a homogeneous random system of 1000 particles in a box of size $10 \times 10 \times h$ with a system height $h = 9$ (*left graph*) and 5 (*right graph*). The *red rectangles* denote results for $R = 0$, the *green circles* for a maximal error of 10^{-2}, corresponding to $R = 1$ for $h = 9$ and $R = 0.1$ for $h = 5$, and the *blue triangles* for a maximal error of 10^{-4}, corresponding to $R = 1.8$ and $R = 0.3$. *Open symbols* show the maximal error that occurred within all particles with similar z-coordinates; the *filled symbols* show the rms force error of these particles. The P^3M method was used as the three-dimensional method, tuned to a rms force error below 10^{-4}. Using the 10^{-4} error bound for ELC, the ELC errors are completely hidden by the P^3M errors, rendering this setting safe for production

veloped earlier can also be used (see however [51]). Again, both are $\mathcal{O}(N^2)$ methods which converge rather slowly, and no error estimates have been provided.

We recently developed a method which we called MMM1D [69], which is basically the MMM2D method adapted to 1D, and will be presented below. Again, rigorous error bounds can be derived, and although the method is $\mathcal{O}(N^2)$, it is much faster than EW1D, and easier to use. We have already applied this method to charged rod systems with good results [65, 66].

As mentioned in Sect. 2.5, in principle the FMM, multigrid methods, or tree codes can handle this situation, but they are much too slow for the normally only small number of charges involved, and error estimates are not easy to obtain. Also, a modified Ewald method in which the summation of the reciprocal-space vectors was modified [70], similar to the one used by Kawata and Mikami [71] exists, but also here the approximations made seem hard to control which render the method rather useless.

And finally, one can in principle again use 3D Ewald methods and separate the rod images sufficiently to ensure that the interactions are negligible, which can be ensured for systems where mobile counterions are present which can screen the interactions. This has been applied often by Deserno et al. [23, 72, 73]. As a side remark, similar to the ELC method, by changing the summation order and using the prescription of Smith [50], a corrected 3D Ewald summation should also be possible, but it might again not be easy to solve the problem of how to control the error.

4.1 MMM1D

MMM1D, as MMM2D and ELC, relies on a screening convergence factor approach. Similar to the two-dimensional case, it can be proven that this approach is equivalent to a spherical summation approach, which in a one-dimensional periodic system reduces to simply adding up in pairs of increasing distance. Due to the symmetry of the problem, only two coordinates are important in the force calculation, namely, the component in the periodic axis, named z, and the component perpendicular to that, termed ρ here. In the spirit of the notation for the other MMM-type methods, we define $\omega = 2\pi u_z p$. For $\rho > 0$ the far formula in one dimension is then given by

$$\begin{aligned}
\phi(\rho, z) &= 4u_z \sum_{p\neq 0} \mathrm{K}_0(\omega\rho)\cos(\omega z) - 2u_z \log\left(\frac{\rho}{2\lambda_z}\right) - 2u_z\gamma \qquad (74)\\
F_\rho(\rho, z) &= 8\pi u_z^2 \sum_{p\neq 0} p\mathrm{K}_1(\omega\rho)\cos(\omega z) + \frac{2u_z}{\rho}\\
F_z(\rho, z) &= 8\pi u_z^2 \sum_{p\neq 0} p\mathrm{K}_0(\omega\rho)\sin(\omega z)\,.
\end{aligned}$$

In analogy to Eq. 48, the total energy is obtained as

$$E = \frac{1}{2}\sum_{i,j} q_i q_j \phi(\rho_{ij}, z_{ij})\,. \qquad (75)$$

Similarly, the force exerted on particle i is given by

$$F_i = q_i \sum_j q_j F(\rho_{ij}, z_{ij})\,, \qquad (76)$$

where $F(\rho, z) = (F_\rho(\rho, z)x/\rho, F_\rho(\rho, z)y/\rho, F_z(\rho, z))$.

For small ρ we obtain the near formulas in one dimension

$$\begin{aligned}
\tilde{\phi}(\rho, z) &= -u_z \sum_{n\geq 0} \binom{-\frac{1}{2}}{n} \frac{\tilde{\psi}^{2n}(N_\psi, u_z z)}{(2n)!}(u_z\rho)^{2n} \qquad (77)\\
&\quad + \sum_{k=1}^{N_\psi - 1}\left(\frac{1}{r_k} + \frac{1}{r_{-k}}\right) - 2u_z\gamma\\
\widetilde{F}_\rho(\rho, z) &= -u_z^3 \sum_{n\geq 0} \binom{-\frac{1}{2}}{n} \frac{\tilde{\psi}^{2n}(N_\psi, u_z z)}{(2n)!}(u_z\rho)^{2n-1}\\
&\quad + \sum_{k=1}^{N_\psi - 1}\left(\frac{\rho}{r_k^3} + \frac{\rho}{r_{-k}^3}\right)
\end{aligned}$$

$$\widetilde{F}_z(\rho,z) = -u_z^2 \sum_{n\geq 0} \binom{-\frac{1}{2}}{n} \frac{\tilde{\psi}^{2n+1}(N_\psi, u_z z)}{(2n)!} (u_z\rho)^{2n} + \sum_{k=1}^{N_\psi-1} \left(\frac{z+k\lambda_z}{r_k^3} + \frac{z-k\lambda_z}{r_{-k}^3} \right),$$

where $\tilde{\psi}^m(N,z) := \psi^m(N_\psi + u_z z) + \psi^m(N_\psi - u_z z)$, and the tilde denotes that the interactions in the primary simulation box are left out and have to be added separately as usual. Otherwise the expressions of the near and far formulas can be exchanged freely. However, the calculation cannot be improved, and one is stuck with the simple $\mathcal{O}(N^2)$ loops implied by Eq. 76. Note that the functions $\tilde{\psi}^m$ are not only defined for write-up convenience: the functions $\tilde{\psi}^m$ have much better convergent Taylor series expansions compared to those of the pure polygamma functions, and therefore it is highly advisable to implement the $\tilde{\psi}^m$ evaluation rather than the polygamma functions in a fast simulation code.

An upper bound on the absolute errors of the force of the far formula for the summation up to, but not including, the finite cutoff P is

$$|\Delta F|_\infty \leq 8\pi u_z^2 \mathrm{K}_1(2\pi u_z P \rho_{\min}) \frac{\mathrm{e}^{2\pi u_z \rho_{\min}}}{2\pi u_z \rho_{\min}} \left(P - 1 + \frac{1}{2\pi u_z \rho_{\min}} \right)$$

provided that $\rho \geq \rho_{\min}$. The $\rho_{\min}$ will in general be the radius where the calculation switches from the far to the near formula.

The near formula can be treated by the Leibniz criterion as usual and no further error analysis is needed. Note that for simplicity of implementation one will normally choose a fixed N_ψ. Moreover, for $(u_z\rho)/(N_\psi - 1/2) \to 1$ the convergence of the polygamma series will become poor. Therefore, one must check that $\rho < \theta\lambda_z(N_\psi - 1/2)$, where $0 < \theta < 1$ is a constant which determines the number of polygamma terms calculated. The larger θ is, the larger ρ may be, but more polygamma terms have to be calculated in the worst case. Usually it is sufficient to fix N_ψ to a value of 2 to avoid the additional summation of direct interactions, which gives $\theta = 2/3$. Numerical tests have shown that the optimal computation time is achieved for $\rho \ll \lambda_z$ in general, therefore this has no impact on the performance of the method.

The error estimates are the same as for MMM, MMM2D, and ELC, so it is unnecessary to show the error dependencies here again. The same warning as for the other methods holds: the error drops exponentially outside the switching radius. On the other hand, since the formulas are radially invariant, the radial error distribution is uniform.

Although MMM1D has a clearly unfavorable scaling of $\mathcal{O}(N^2)$, it can still be used effectively in a simulation, especially because of its small prefactors.

For the calculation of the forces for 400 charged particles with a precision of 10^{-4}, the MMM1D method can be implemented such that the force calculation takes less than 10 ms per time step on a 2-GHz PC.

5 Maggswellian Dynamics

There is a recent development for treating electrostatic interactions on and off lattice, introduced first by A.C. Maggs' group, and later by B. Dünweg and I. Pasichnyk [21, 74, 75]. Maggs used something similar to the Maxwell equations on a lattice. After all, electrodynamics is a local field theory, so one should be able to implement the equations locally on a lattice, which then results in an algorithm that must scale linearly in N. The algorithm, due to its locality, is especially suited for Monte Carlo (MC) algorithms [21]. In that version, the pseudo magnetic field is propagated via a diffusive process through the lattice, and MC moves are constructed that do not violate the Gauss law. Maggs showed that the diffusive dynamics recovers all static properties in thermal equilibrium, i.e., for long times. The whole formalism, where the charges reside on lattice sites, the electric and magnetic fields are link variables, and the electromagnetic field tensor analogon is a plaquette variable, is similar in spirit to the one used in lattice gauge theories [76]. A particularly nice feature, which so far has not been exploited, is the possibility to assign a locally varying dielectric constant. This is a problem which puts a formidable constraint on most other methods.

Molecular dynamics variants of this method were presented in Refs. [74, 75]. Here, the $1/r$ interaction is represented by a local interaction that propagates in waves with a tunable speed of light, c. It is sufficient to make c small enough such that the quasistatic approximation is still valid. This is, in spirit, like performing a Car–Parrinello [77] simulation, where the electrons are assigned an unphysically large mass in order to slow down those degrees of freedom. Again, the static properties of the system do not depend on the value chosen for c. This, and the lattice spacing, are the principal parameters that have to be tuned for speed and accuracy. Preliminary benchmarks presented in Ref. [74] show that the so-called MEMD (Maxwell equation molecular dynamics) method has a similar performance to our implementation of the P^3M algorithm. The possible advantages of MEMD are that it should be easier to parallelize and can be used for systems with locally varying dielectric constants, but more results are needed to be able to assess all the advantages and disadvantages compared to the previously presented methods.

6 The Dipolar Ewald Summation in 3D

In the previous section we treated methods to deal with the Coulomb interaction or, more generally, interactions which vary with $1/r$. However, there are also systems of interest, which can be modeled by interacting point dipoles. Substances of that kind are ferrofluids, which are basically dispersed magnetic particles [78], magnetorheological (MR) or electrorheological (ER) fluids, or solvents which can be modeled approximately by dipolar interactions like water (see the contribution of J.-J. Weis and D. Levesque in this volume). For periodic geometries again the Ewald method can be used. In the simplest implementations the involved computation time grows like N^2, or at best like $N^{3/2}$, if the cutoff is optimally varied with the splitting parameter [9]. For MC simulations, knowing the energy formulas is sufficient, whereas for MD simulations, we need to know forces and torques. In this section we give a reliable error estimate for the energy, forces, and torques, when computed via the standard Ewald sum. Moreover, we present a detailed discussion on the optimization of the parameters, which will lead to the most efficient parameters for a predefined error in each observable quantity. This can all be done prior to the actual simulation, thus ensuring optimal performance at optimally controlled errors. The exposition is heavily based on the paper by Wang and Holm [79].

Throughout this section we return to the notations already used in Sect. 2.1 in the discussion of the Ewald-type methods for point charges. Consider a system of N particles with a point dipole $\boldsymbol{\mu}_i$ at their center position $\boldsymbol{r}_i$ in a cubic simulation box of length L. If periodic boundary conditions are applied, the total electrostatic energy of the box is given by

$$U = \frac{1}{2} \sum_{\boldsymbol{n} \in \mathbb{Z}^3} \sum_{i,j=1}^{N}{}' \left\{ \frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j}{| \boldsymbol{r}_{ij} + \boldsymbol{n} |^3} - \frac{3[\boldsymbol{\mu}_i \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})][\boldsymbol{\mu}_j \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})]}{| \boldsymbol{r}_{ij} + \boldsymbol{n} |^5} \right\}, \tag{78}$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$. The sum over $\boldsymbol{n}$ is over all simple cubic lattice points, $\boldsymbol{n} = (n_x L, n_y L, n_z L)$, with n_x, n_y, n_z integers. The prime indicates that the $i = j$ term must be omitted for $\boldsymbol{n} = 0$. The slowly decaying long-range part of the dipolar potential makes the straightforward summation of Eq. 78 too time-consuming. The Ewald trick splits the problem again into two rapidly convergent parts, one in real space and one in reciprocal space. The details of the method are discussed in Refs. [5, 7, 9, 29]. Here we only give the final expressions

$$U = U^{(r)} + U^{(k)} + U^{(s)} + U^{(d)}, \tag{79}$$

where the real-space $U^{(r)}$, the k-space (reciprocal-space) $U^{(k)}$, the self $U^{(s)}$, and the dipole (surface) $U^{(d)}$ contributions are respectively given by:

$$U^{(r)} = \frac{1}{2} \sum_{\boldsymbol{n}\in\mathbb{Z}^3} \sum_{i,j=1}^{N}{}' \Big\{ (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j) B(|\, \boldsymbol{r}_{ij} + \boldsymbol{n} \,|) \\ - [\boldsymbol{\mu}_i \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})][\boldsymbol{\mu}_j \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})] C(|\, \boldsymbol{r}_{ij} + \boldsymbol{n} \,|) \Big\}, \tag{80}$$

$$U^{(k)} = \frac{1}{2L^3} \sum_{\boldsymbol{k}\in\mathbb{Z}^3, \boldsymbol{k}\neq 0} \frac{4\pi}{k^2} \mathrm{e}^{-(\pi k/\alpha L)^2} \sum_{i,j=1}^{N} (\boldsymbol{\mu}_i \cdot \boldsymbol{k})(\boldsymbol{\mu}_j \cdot \boldsymbol{k}) \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}\cdot\boldsymbol{r}_{ij}/L}, \tag{81}$$

$$U^{(s)} = -\frac{2\alpha^3}{3\sqrt{\pi}} \sum_{i=1}^{N} \mu_i^2, \tag{82}$$

$$U^{(d)} = \frac{2\pi}{(2\varepsilon' + 1)L^3} \sum_{i,j=1}^{N} \boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j. \tag{83}$$

The sums over i and j are for the particles in the central box and

$$B(r) = [\mathrm{erfc}(\alpha r) + (2\alpha r/\sqrt{\pi}) \exp(-\alpha^2 r^2)]/r^3, \tag{84}$$

$$C(r) = [3\mathrm{erfc}(\alpha r) + (2\alpha r/\sqrt{\pi})(3 + 2\alpha^2 r^2) \exp(-\alpha^2 r^2)]/r^5. \tag{85}$$

The inverse length α is the splitting parameter of the Ewald summation which should be chosen so as to optimize the performance. The form Eq. 83 given for the surface correction assumes that the set of the periodic replications of the simulation box tends in a spherical way toward an infinite cluster, and that the medium outside this sphere is a uniform dielectric with dielectric constant ε' [7, 29]. The case of a surrounding vacuum corresponds to $\varepsilon' = 1$ and the surface term vanishes for the metallic boundary conditions ($\varepsilon' = \infty$).

In practical calculations, the infinite sums in Eqs. 80 and 81 are truncated by only taking into account distances which are smaller than some real-space cutoff r_c and wave vectors with a modulus smaller than some reciprocal space cutoff k_c. If $r_c \leq L/2$, the sum in real space (Eq. 80) reduces to the normal minimum image convention. The double sum over particles in $U^{(k)}$ can be replaced by a product of two single sums, which is more suitable for numerical calculations.

The force $\boldsymbol{F}_i$ acting on particle i is obtained by differentiating the potential energy U with respect to $\boldsymbol{r}_i$, i.e.,

$$\boldsymbol{F}_i = -\frac{\partial U}{\partial \boldsymbol{r}_i} = \boldsymbol{F}_i^{(r)} + \boldsymbol{F}_i^{(k)}, \tag{86}$$

with the real-space and k-space contributions given by:

$$\boldsymbol{F}_i^{(r)} = \sum_{\boldsymbol{n}\in\mathbb{Z}^3} \sum_{j=1}^{N}{}' \Big\{ \Big\{ (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j)(\boldsymbol{r}_{ij} + \boldsymbol{n}) + \boldsymbol{\mu}_i [\boldsymbol{\mu}_j \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})] + [\boldsymbol{\mu}_i \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})] \boldsymbol{\mu}_j \Big\} C(|\, \boldsymbol{r}_{ij} + \boldsymbol{n} \,|) - [\boldsymbol{\mu}_i \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})][\boldsymbol{\mu}_j \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})] D(|\, \boldsymbol{r}_{ij} + \boldsymbol{n} \,|)(\boldsymbol{r}_{ij} + \boldsymbol{n}) \Big\}, \tag{87}$$

$$\boldsymbol{F}_i^{(k)} = \frac{2\pi}{L^4} \sum_{j=1}^{N} \sum_{\boldsymbol{k}\in\mathbb{Z}^3, \boldsymbol{k}\neq 0} \frac{4\pi \boldsymbol{k}}{k^2} \Big[(\boldsymbol{\mu}_i \cdot \boldsymbol{k})(\boldsymbol{\mu}_j \cdot \boldsymbol{k}) \exp[-(\pi k/\alpha L)^2] \sin(2\pi \boldsymbol{k} \cdot \boldsymbol{r}_{ij}/L) \Big], \tag{88}$$

where

$$D(r) = [15\mathrm{erfc}(\alpha r) + (2\alpha r/\sqrt{\pi})(15 + 10\alpha^2 r^2 + 4\alpha^4 r^4) \exp(-\alpha^2 r^2)]/r^7 \,. \tag{89}$$

Since the self- and surface-energy terms (Eqs. 82, 83) are independent of the particle positions, they have no contributions to the force, unlike the Ewald summation for the Coulomb systems where the surface term contributes.

The torque $\boldsymbol{\tau}_i$ acting on particle i is related to the electrostatic field $\boldsymbol{E}_i$ at the location of this particle,

$$\boldsymbol{\tau}_i = \boldsymbol{\mu}_i \cdot \boldsymbol{E}_i = \boldsymbol{\tau}_i^{(r)} + \boldsymbol{\tau}_i^{(k)} + \boldsymbol{\tau}_i^{(d)} \,, \tag{90}$$

with

$$\boldsymbol{E}_i = -\frac{\partial U}{\partial \boldsymbol{\mu}_i} \,, \tag{91}$$

and thus

$$\boldsymbol{\tau}_i^{(r)} = -\sum_{\boldsymbol{n}\in\mathbb{Z}^3} \sum_{j=1}^{N}{}' \Big\{ (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j) B(|\, \boldsymbol{r}_{ij} + \boldsymbol{n} \,|) - [\boldsymbol{\mu}_i \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})][\boldsymbol{\mu}_j \cdot (\boldsymbol{r}_{ij} + \boldsymbol{n})] C(|\, \boldsymbol{r}_{ij} + \boldsymbol{n} \,|) \Big\}, \tag{92}$$

$$\boldsymbol{\tau}_i^{(k)} = -\frac{1}{L^3} \sum_{j=1}^{N} \sum_{\boldsymbol{k}\in\mathbb{Z}^3, \boldsymbol{k}\neq 0} \frac{4\pi}{k^2} (\boldsymbol{\mu}_i \cdot \boldsymbol{k})(\boldsymbol{\mu}_j \cdot \boldsymbol{k}) \, \mathrm{e}^{-(\pi k/\alpha L)^2} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k} \cdot \boldsymbol{r}_{ij}/L} \,, \tag{93}$$

$$\boldsymbol{\tau}_i^{(d)} = -\frac{4\pi}{(2\varepsilon' + 1)L^3} \sum_{j=1}^{N} \boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j \,. \tag{94}$$

6.1 Error Formulas

We now give estimates for the rms error caused by cutting off the Ewald summation in real space and k-space for the forces, total energy, and torques. There are no errors involved in the self- and surface contributions (Eqs. 82, 83, 94), because no cutoff operations are applied to them.

As can be shown similarly to the method of presentation in Sect. 2.4, the rms error for the force can be cast in the following form

$$\Delta F^{(r)} \approx \chi^{(r)} \frac{\mathcal{M}^2}{\sqrt{N}}, \tag{95}$$

where the quantity $\mathcal{M}$ is defined as

$$\mathcal{M}^2 := \sum_{j=1}^{N} \mu_j^2 . \tag{96}$$

After some lengthy calculations one obtains

$$\chi^{(r)2} \approx L^{-3} r_c^{-9} \alpha^{-4} \left(\frac{13}{6} C_c^2 + \frac{2}{15} D_c^2 - \frac{13}{15} C_c D_c\right) \exp(-2\alpha^2 r_c^2), \tag{97}$$

where the terms C_c and D_c are given by

$$C_c = 4\alpha^4 r_c^4 + 6\alpha^2 r_c^2 + 3, \tag{98}$$

$$D_c = 8\alpha^6 r_c^6 + 20\alpha^4 r_c^4 + 30\alpha^2 r_c^2 + 15. \tag{99}$$

The resulting rms expectation of the real-space cutoff error in the forces is thus

$$\Delta F^{(r)} \approx \mathcal{M}^2 (L^3 \alpha^4 r_c^9 N)^{-1/2} \left(\frac{13}{6} C_c^2 + \frac{2}{15} D_c^2 - \frac{13}{15} C_c D_c\right)^{1/2} \exp(-\alpha^2 r_c^2). \tag{100}$$

The derivation of the expected real-space cutoff errors in the total potential energy and torques follows the same path. For calculating the fluctuation of the error in total energy, it is noted that the interaction energy between two dipoles is evenly shared between them. That means the sum of $\langle(\Delta U^{(r)})^2\rangle$ over all particles contains each pair contribution twice, and thus the fluctuation of the real-space cutoff error is one half of the sum [35]. Then the rms value of the real-space cutoff error of the total potential energy is estimated as

$$\Delta U^{(r)} \approx \mathcal{M}^2 (L^3 \alpha^4 r_c^7 N)^{-1/2} \left[\frac{1}{4} B_c^2 + \frac{1}{15} C_c^2 - \frac{1}{6} B_c C_c\right]^{1/2} \exp(-\alpha^2 r_c^2) \tag{101}$$

with

$$B_c = 2\alpha^2 r_c^2 + 1. \tag{102}$$

The rms error on the torques is estimated similarly to the force as

$$\Delta\tau^{(r)} \approx \mathcal{M}^2(L^3\alpha^4 r_c^7 N)^{-1/2}\left[\frac{1}{2}B_c^2 + \frac{1}{5}C_c^2\right]^{1/2}\exp(-\alpha^2 r_c^2)\,. \tag{103}$$

Equations 100, 101, and 103 all contain the exponential $\exp(-\alpha^2 r_c^2)$. For sufficiently low errors, αr_c has to be larger than one, for example $\alpha r_c \approx \pi$ for an error of $\exp(-\pi^2) \approx 5\times 10^{-5}$. If only the highest powers of αr_c are retained, the estimates of the real-space cutoff errors in the total energy, forces, and torques can be reduced to

$$\Delta U^{(r)} \approx 4\mathcal{M}^2\alpha^2(r_c/15NL^3)^{1/2}\exp(-\alpha^2 r_c^2)\,, \tag{104}$$

$$\Delta F^{(r)} \approx 8\mathcal{M}^2\alpha^4(2r_c^3/15NL^3)^{1/2}\exp(-\alpha^2 r_c^2)\,, \tag{105}$$

$$\Delta\tau^{(r)} \approx 4\mathcal{M}^2\alpha^2(r_c/5NL^3)^{1/2}\exp(-\alpha^2 r_c^2)\,, \tag{106}$$

where Eq. 104 is a factor of $\sqrt{6/5}$ slightly larger than that given in Eq. 35 of Ref. [35]. The advantage of these simplified formulas is that they reflect the dependence of the rms errors on α and r_c more directly, and thus could be used more easily in determining the optimal values of these parameters.

In deriving the estimates of the reciprocal-space (k-space) cutoff errors, we further assume that the radial distribution function of the particles is approximately unity at all distances. Following Eq. 88, the k-space cutoff error in the force acting on particle i is given by

$$\Delta\boldsymbol{F}_i^{(k)} = \sum_{j=1}^{N}\sum_{\boldsymbol{k},k>k_c}\frac{8\pi^2\boldsymbol{k}}{L^4k^2}(\boldsymbol{\mu}_i\cdot\boldsymbol{k})(\boldsymbol{\mu}_j\cdot\boldsymbol{k}) \times\exp[-(\pi k/\alpha L)^2]\sin(2\pi\boldsymbol{k}\cdot\boldsymbol{r}_{ij}/L)\,. \tag{107}$$

Note that the diagonal term ($j = i$) in the sum does not depend on the positions of the particles. It will provide a systematic contribution to the cutoff error in k-space [35]. In Eq. 107 this contribution equals zero, thus there is no systematic part of the error in the forces. The same thing happens to the error in the torques. But for the total energy the diagonal terms are positive and the systematic contribution plays a dominant role in the cutoff error.

The off-diagonal terms in Eq. 107 do depend on the positions of the particles and have alternating signs. Similarly to before, the off-diagonal contribution to the cutoff error in $\Delta\boldsymbol{F}_i^{(k)}$ is given by

$$\Delta\boldsymbol{F}_{i,\mathrm{off}}^{(k)} = \mid\boldsymbol{\mu}_i\mid\sum_{j\neq i}\mid\boldsymbol{\mu}_j\mid\boldsymbol{\chi}_{ij}^{(k)} \tag{108}$$

with

$$\boldsymbol{\chi}_{ij}^{(k)} = \sum_{\boldsymbol{k},k>k_c} \frac{8\pi^2\boldsymbol{k}}{L^4}\Big(\cos\vartheta(\hat{\boldsymbol{\mu}}_i,\hat{\boldsymbol{k}})\cos\vartheta(\hat{\boldsymbol{\mu}}_j,\hat{\boldsymbol{k}})\exp[-(\pi k/\alpha L)^2] \cdot \mathrm{i}\exp(2\pi\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}/L)\Big)\,, \tag{109}$$

where $\boldsymbol{r}$ stands for $\boldsymbol{r}_{ij}$ and $\sin(2\pi\boldsymbol{k}\cdot\boldsymbol{r}/L)$ is rewritten as $\mathrm{i}\exp(2\pi\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}/L)$ according to the symmetrical character of the summation over $\boldsymbol{k}$. $\hat{\boldsymbol{k}}$ is the unit vector along $\boldsymbol{k}$. Since the particles are assumed to be randomly distributed over the simulation box, the fluctuation of $\Delta\boldsymbol{F}_{i,\mathrm{off}}^{(k)}$ can also be written as

$$\left\langle(\Delta\boldsymbol{F}_{i,\mathrm{off}}^{(k)})^2\right\rangle = \mu_i^2\sum_{j\neq i}\sum_{m\neq i}|\,\boldsymbol{\mu}_j\,||\,\boldsymbol{\mu}_m\,|\,\langle\boldsymbol{\chi}_{ij}^{(k)}\cdot\boldsymbol{\chi}_{im}^{(k)*}\rangle \approx \mu_i^2\mathcal{M}^2\chi^{(k)2}\,. \tag{110}$$

Again $\chi^{(k)2}$ is independent of i and j.

Choosing the z-axis of the spherical coordinates (k,θ,ϕ) along the $\hat{\boldsymbol{\mu}}$ orientation, we find

$$\chi^{(k)2} \approx 128\pi^3L^{-6}\alpha^2k_c^3\exp[-2(\pi k_c/\alpha L)^2]/15\,. \tag{111}$$

The rms expectation of the k-space cutoff error in the forces is thus

$$\Delta F^{(k)} \approx 8\pi\mathcal{M}^2L^{-3}\alpha(2\pi k_c^3/15N)^{1/2}\exp[-(\pi k_c/\alpha L)^2]\,. \tag{112}$$

Here the notation $\Delta F_{\mathrm{off}}^{(k)}$ is replaced directly with $\Delta F^{(k)}$ due to the fact of no diagonal contribution.

The derivation of the off-diagonal parts of the cutoff errors in the total energy and torques proceeds in the same way. That the sum over $\langle(\Delta U_{i,\mathrm{off}}^{(k)})^2\rangle$ contains each pair contribution twice has also been considered in the error estimate of the total energy. The results are given by

$$\Delta U_{\mathrm{off}}^{(k)} \approx 4\mathcal{M}^2L^{-2}\alpha(\pi k_c/15N)^{1/2}\exp[-(\pi k_c/\alpha L)^2]\,, \tag{113}$$

$$\Delta\tau^{(k)} \approx 4\mathcal{M}^2L^{-2}\alpha(\pi k_c/5N)^{1/2}\exp[-(\pi k_c/\alpha L)^2]\,. \tag{114}$$

$\Delta\tau^{(k)}$ is also used directly instead of $\Delta\tau_{\mathrm{off}}^{(k)}$.

The diagonal (systematic) part of the cutoff error in the total energy can be written as

$$\begin{aligned}\Delta U_{\mathrm{diag}}^{(k)} &= \frac{1}{2\sqrt{N}}\sum_{i=1}^{N}\sum_{\boldsymbol{k},k>k_c}\frac{4\pi}{L^3}\mu_i^2\cos^2\vartheta(\hat{\boldsymbol{\mu}}_i,\hat{\boldsymbol{k}})\exp[-(\pi k/\alpha L)^2] \\ &\approx \frac{2\pi}{L^3\sqrt{N}}\mathcal{M}^2\int_{k_c}^{\infty}\exp[-(\pi k/\alpha L)^2]k^2\mathrm{d}k\int_0^{\pi}\sin\theta\mathrm{d}\theta\int_0^{2\pi}\cos^2\vartheta(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{k}})\mathrm{d}\phi \\ &\approx \frac{4}{3}\mathcal{M}^2L^{-1}\alpha^2k_cN^{-1/2}\exp[-(\pi k_c/\alpha L)^2]\,,\end{aligned} \tag{115}$$

where the sum is again approximated by an integral, and an integral expansion formula [35] is used to get the final estimate. The total k-space cutoff error in the total energy is thus

$$\Delta U^{(k)} = \Delta U^{(k)}_{\text{diag}} + \Delta U^{(k)}_{\text{off}} . \tag{116}$$

Comparing Eq. 115 with Eq. 113, it can be seen that the systematic part of the error is a factor of $\sim L\alpha k_c^{1/2} (\gg 1)$ larger than the statistical part. Hence, the systematic contribution is dominant in the k-space cutoff error of the total energy.

Assuming that the real-space and reciprocal-space contributions to the error are independent, the total cutoff error in Ewald summation can be written as

$$\Delta\Theta = \sqrt{\Delta\Theta^{(r)2} + \Delta\Theta^{(k)2}} , \tag{117}$$

where Θ stands for U, F, and τ.

6.2 Optimization of Parameters

In this section, we discuss the use of the analytical formulas derived in Sect. 3 to determine the optimal values of α, r_c, and k_c by which the required accuracy could be satisfied and the computation time minimized. The detailed discussions on this subject can also be found in Refs. [9, 46].

The overall computation time for computing the forces with the Ewald method is approximately given by [46]

$$\mathcal{T} = a_r N^2 (r_c/L)^3 + a_k N k_c^3 , \tag{118}$$

where the primitive overheads a_r and a_k highly depend on the implementation of the code and need to be found by numerical experiments. As an example, we have carried out the time experiments on a DEC personal workstation (CPU 433 MHz) using a standard Fortran 77 compiler. In the implementation the complementary error function and its derivative are calculated with table lookup and the reciprocal-space summation is optimized as in Refs. [9, 29]. The linked-cell method is used to deal with the short-range forces (when doing simulations). The primitive overheads are then found roughly to be $a_r = 2.5\ \mu$s and $a_k = 0.7\ \mu$s.

For a required accuracy δ, the parameters α, r_c, and k_c should be chosen to minimize $\mathcal{T}$ with respect to the two constraints of the error bounds Eqs. 100

and 112, which are restated as

$$\frac{\delta}{\sqrt{2}} = \mathcal{M}^2 \left(L^3\alpha^4 r_c^9 N\right)^{-1/2} \left(\frac{13}{6}C_c^2 + \frac{2}{15}D_c^2 - \frac{13}{15}C_c D_c\right)^{1/2} \exp(-\alpha^2 r_c^2)\,, \quad (119)$$

$$\frac{\delta}{\sqrt{2}} = 8\pi\mathcal{M}^2 L^{-3}\alpha(2\pi k_c^3/15N)^{1/2} \exp[-(\pi k_c/\alpha L)^2]\,. \quad (120)$$

In the case of $\delta \leq 5\times 10^{-5}$, Eq. 105 could be used instead of Eq. 100 to show the dependence of the accuracy on the parameters more clearly. Equations 119 and 120 provide the qualitative function relations of r_c and k_c with α as: $r_c(\alpha) \approx -A\sqrt{\ln\delta}/\alpha$ and $k_c(\alpha) \approx -B\sqrt{\ln\delta}\alpha$. Inserting them into Eq. 118 and differentiating it with respect to α yields $\alpha \propto N^{1/6}$, and thus $r_c \propto N^{-1/6}$ and $k_c \propto N^{1/6}$. The minimized computation time is then proportional to $N^{3/2}$ with the proportionality constant depending on the accuracy. The same results can be found for the Coulomb Ewald method in Refs. [9, 35, 46]. This can be easily understood by comparing Eqs. 100 and 112 in Sect. 3 of this paper with Eqs. 18 and 32 in Ref. [35], and finding the same exponential dependences of the cutoff errors on α, r_c, and k_c for the dipolar and Coulomb Ewald summations.

The numerical investigation of the functional dependence of the optimal parameters on N and δ are performed by using the primitive overheads obtained above. For each given N and δ, we at first choose different values for r_c within the inequality $r_c \leq L/2$. For each r_c the parameters α and k_c are calculated by solving Eqs. 119 and 120. These values are then introduced into Eq. 118 to figure out the optimal value of r_c which gives the minimum computation time. In calculations the size of the simulation cell is fixed to a dimensionless length of $L = 10$. The range of accuracy requirement and number of particles are chosen to be $\delta = 10^{-2}$ to 10^{-5} measuring in $\mathcal{P}^2/\mathcal{L}^4$ and $N = 10^3$ to $N = 10^6$, which should cover most of the applications. The particles are supposed to have a uniform dipole moment of $\mathcal{P}$. The results for the optimal values of the parameters and the corresponding computation time per particle are shown in Fig. 10a–d, respectively. It can be clearly seen that the functional dependence of the parameters and the overall computation time on N are just as discussed above. Figure 10c shows that when a high accuracy is required for a system with a small number of particles, the predicted real-space cutoff is larger than half of the box length and $r_c = L/2$ must be used. The optimal α values hardly depend on the accuracies. These results are very similar to that obtained for the Coulomb Ewald summation in Ref. [46], except for $r_c \propto N^{-1/6}$ here and $r_c \propto N^{1/6}$ there. This is because Petersen considered a system of constant density, while we chose the volume of the simulation cell to be constant.

Finally we should remark that besides the standard Ewald method, dipolar variants of basically all Coulomb methods are possible. Known worked-out examples can be found in the contribution of Weis and Levesque in this vol-

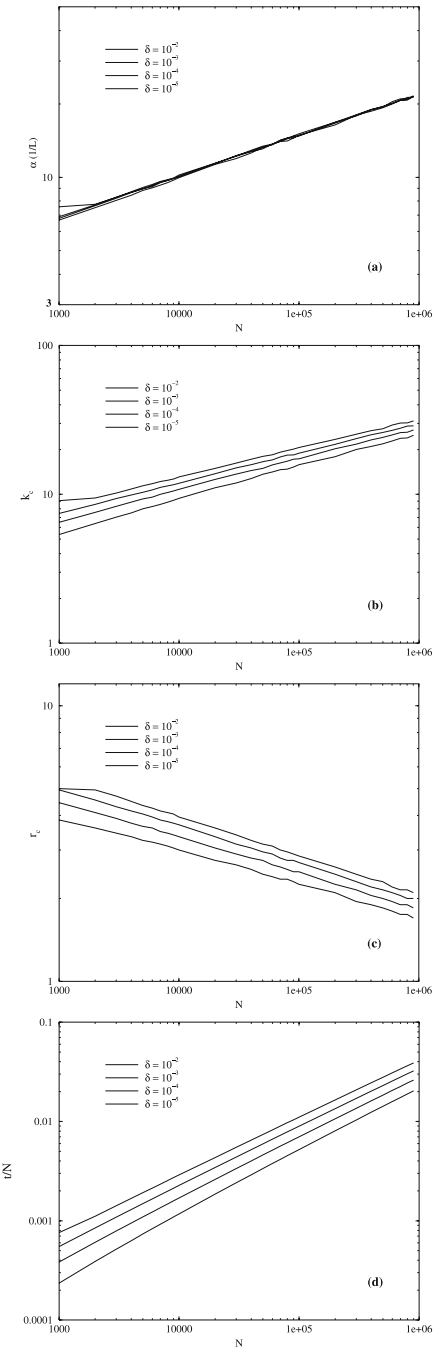

Fig. 10 Optimal values of the parameters α (**a**), k_c (**b**) , and r_c (**c**) as well as the corresponding minimized computation time $\mathcal{T}/N$ (**d**) as a function of the number of particles

ume. Very recently, the ELC term was also derived for dipolar systems [80], and a dipolar SPME has been described in Ref. [81]. We have recently also derived the optimal influence function for the dipolar P^3M and the error formulas for the dipolar cases, which again should yield the optimal mesh-based method for dipolar systems. In general, most of our previous discussions related to the efficiency and applicability of the Coulomb methods should be relevant to the dipolar cases as well.

7 ESPResSo

In the following discussion we will briefly describe a simulation package that contains P^3M and MEMD [74] for electrostatic interactions in 3D periodic boundary conditions, MMM2D and ELC for two periodic dimensions, and MMM1D for only one periodic dimension, and the dipolar Ewald sum. This program package is called ESPResSo standing for **E**xtensible **S**imulation **P**ackage for **Res**earch on **So**ft matter [3, 4].

The main goal of ESPResSo was to have a simulation package that contains state-of-the-art algorithms for simulations of bead–spring models in a parallel code that is extensible. The extensibility paradigm requires that readability of the code is preferred over special code optimization in ESPResSo and that for many basic tasks standardized interfaces exist. The code kernel is written in ANSI C, and runs on a variety of hardware platforms like PCs (GNU/Linux on IA32 and AMD64 processors), workstations (MacOS on PowerPC processors and Tru64 on Alpha processors), and high-performance servers (AIX on Power4 processors).

The wide field of simulation topics investigated by ESPResSo requires a high flexibility of the simulation code, which in ESPResSo is obtained by a script language which is used to control the simulation process. The simulation control script determines all simulation parameters such as the number and type of particles, the type of interactions between these particles, and how the system is propagated. The script language used in these scripts allows one to change most of the parameters even during the simulation. This flexibility makes it possible to perform highly complex simulation procedures, such as adapting the interaction parameters to the current configuration during the simulation, cooling down the system in a simulated annealing process, or applying or removing a constraint.

ESPResSo is not a self-contained code, but relies on other open-source packages. Most prominent is the use of the Tcl [82] script language interpreter for simulation control. For the parallelization standard MPI routines are used, which on Linux and MacOS are provided, for example, by the LAM/MPI [83] implementation, or MPICH [84]. The P^3M method for the electrostatic in-

teraction and the mode analysis tool for membranes rely on the FFTW [85]. Besides these libraries, which are required to be able to have a running version of ESPResSo, the development process is supported heavily by the use of the CVS version control system [86], which allows a large number of developers to work simultaneously on the code.

Where available, we used the error estimates to construct tuning routines, which automatically put the algorithmic parameters close to their optimal values for the desired accuracies. The source code of this software, and consequently of the methods, is freely available. The exact procedures and current status of the program package can be found on the web at http://www.espresso.mpg.de.

8 Concluding Remarks

We have given a review of methods to compute long-range interactions in fully or, for the case of the Coulomb sum, also partially periodic boundary conditions. We tried to give reasons why one should worry about errors, and how error estimates can be used to tune the algorithms to perform optimally for speed and accuracy. We have also tried to convey our understanding to the reader of when each algorithm should be used.

Most of the algorithms important to the field of charged soft matter have been implemented within our research group into the program package ESPResSo. This program is open source, and information regarding the latest version, or information on how to participate in the further development of this package, can be found at http://www.espresso.mpg.de.

Acknowledgements This article would have been impossible without the contributions of J. Dejoannis, M. Deserno, H.J. Limbach, B.A. Mann, and Z. Wang. Helpful discussions with K. Kremer and other members of our pep group are gratefully acknowledged. This work was supported by the BMBF 03N 6500, and the DFG within the SFB 625, TR6, and grants HO-1108/8-3 and HO-1108/11-2.

References

1. Jones RAL (2002) Soft condensed matter. Oxford University Press, Oxford
2. Holm C, Kékicheff P, Podgornik R (2001) NATO Science Series II: Mathematics, physics and chemistry, vol 46. Kluwer, Dordrecht
3. Arnold A, Mann BA, Limbach H-J, Holm Christian (2004) In: Kremer K, Macho V (eds) Forschung und wissenschaftliches Rechnen 2003, vol 63 of GWDG-Bericht. Gesellschaft für wissenschaftliche Datenverarbeitung mbh, Göttingen, Germany, p 43
4. Arnold A, Mann BA, Limbach H-J, Holm C. ESPResSo – An Extensible Simulation Package for Research on Soft Matter Systems. Comp Phys Comm (in press)

5. Ewald PP (1921) Ann Phys 64:253
6. Heyes DM (1981) J Chem Phys 74:1924
7. de Leeuw SW, Perram JW, Smith ER (1980) Proc R Soc Lond A 373:27
8. de Leeuw SW, Perram JW, Smith ER (1980) Proc R Soc Lond A 373:57
9. Perram J, Petersen GH, de Leeuw S (1988) Mol Phys 65:875
10. Deserno M, Holm C (1998) J Chem Phys 109:7678
11. Lekner J (1991) Physica A 176:485
12. Sperb R (1998) Mol Simulat 20:179
13. Sperb R (1999) Mol Simulat 22:199
14. Strebel R, Sperb R (2001) Mol Simulat 27:61
15. Arnold A (2004) PhD thesis, Johannes Gutenberg-University, Mainz
16. Barnes JE, Hut P (1986) Nature 324:446
17. Greengard L, Rhoklin V (1987) J Comput Phys 73:325
18. Esselink K (1995) Comput Phys Commun 87:375
19. Sagui C, Darden T (2001) J Chem Phys
20. Tsukerman I (2004) IEEE Trans Magn 40:2158
21. Maggs AC, Rosseto V (2002) Phys Rev Lett 88:196402
22. Deserno M, Holm C (1998) J Chem Phys 109:7694
23. Deserno M (2000) PhD thesis, Universität Mainz
24. Arnold A, Holm C (2002) Comput Phys Commun 148:327
25. Arnold A, de Joannis J, Holm C (2002) J Chem Phys 117:2496
26. de Joannis J, Arnold A, Holm C (2002) J Chem Phys 117:2503
27. Attig N, Binder K, Grubmüller H, Kremer K (eds) (2004) Efficient methods for long-range interactions in periodic geometries plus one application. NIC series, vol 23. Research Centre Jülich
28. Frenkel D (2002) Science 296:65
29. Allen MP, Tildesley DJ (1987) Computer simulation of liquids, 1st edn. Oxford Science, Oxford
30. Caillol J-M (1994) J Chem Phys 101:6080
31. Boresch S, Steinhauser O (1997) Ber Bunsenges Phys Chem 101:1019
32. Smith ER (1988) Mol Phys 65:1089
33. Berendsen HJC (1993) In: van Gunsteren WF, Weiner PK, Wilkinson AJ (eds) Computer simulation of biomolecular systems, vol 2. Escom, Leiden, p 161
34. Hünenberger PH (2000) J Chem Phys 113:10464
35. Kolafa J, Perram JW (1992) Mol Simulat 9:351
36. Hummer G, Pratt LR, García AE (1998) J Phys Chem A 102:7885
37. Sangster MJL, Dixon M (1976) Adv Phys 25:247
38. Adams DJ, Dubey GS (1987) J Comput Phys 72:156
39. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical Recipes in C, 2nd edn. Cambridge University Press, Cambridge
40. Hockney RW, Eastwood JW (1988) Computer simulation using particles. IOP, London
41. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089
42. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen L J Chem Phys 103:8577
43. Pollock EL, Glosli J (1996) Comput Phys Commun 95:93
44. Limbach HJ (2001) PhD thesis, Johannes Gutenberg Universität, Mainz
45. Schoenberg IJ (1973) Cardinal Spline Interpolation. Society for Industrial and Applied Mathematics, Philadelphia
46. Petersen HG (1995) J Chem Phys 103:3668

47. Greengard L (1988) The rapid evaluation of potential fields in particle systems. MIT Press, Cambridge, MA
48. Greengard L, Rokhlin V (1997) Acta Numerica 6:229
49. Lekner J (1989) Physica A 157:826
50. Smith ER (1981) Proc R Soc Lond A 375:475
51. Mazars M (2001) J Chem Phys 115:2955
52. Moreira AG, Netz RR (2001) Phys Rev Lett 87:078301
53. Sperb R (1994) Mol Simulat 13:189
54. Strebel R (1999) Dissertation 13504, ETH, Zurich
55. Widmann AH, Adolf DB (1997) Comput Phys Commun 107:167
56. Kawata M, Nagashima U (2001) Chem Phys Lett 340:165
57. Mazars M (2002) J Chem Phys 117:3524
58. Arnold A, Holm C (2002) Chem Phys Lett 354:324
59. Shelley JC, Patey GN (1996) Mol Phys 88:385
60. Spohr E J Chem Phys 107:6342
61. Yeh I-C, Berkowitz ML (1999) J Chem Phys 111:3155
62. Mináry P, Tuckerman ME, Pihakari KA, Martyna GJ (2002) J Chem Phys 116:5351
63. Arnold A (2001) Diploma thesis, Johannes Gutenberg-Universität
64. Deserno M, Holm C, May S (2000) Macromolecules 33:199
65. Deserno M, Arnold A, Holm C (2003) Macromolecules 36:249
66. Naji A, Arnold A, Holm C, Netz RR (2004) Europhys Lett 67:130
67. Porto M (2000) J Phys A 33:6211
68. Langridge DJ, Hart JF, Crampin S (2001) Comput Phys Commun 134:78
69. Arnold A, Holm C MMM1D: A method for calculating electrostatic interactions in 1D periodic geometries. Comp Phys Comm (in press)
70. Bródka A (2002) Chem Phys Lett 363:604
71. Kawata M, Mikami M (2001) Chem Phys Lett 340:157
72. Deserno M, Holm C, Kremer K (2001) Molecular dynamics simulations of the cylindrical cell model. Surfactant science series, vol 99. Marcel Dekker, New York, p 59
73. Deserno M, Holm C (2002) Mol Phys 100:2941
74. Pasichnyk I, Dünweg B (2004) J Phys Condens Mat 16:3999
75. Rottler J, Maggs AC (2004) Phys Rev Lett 93:170201
76. Wilson KG (1974) Phys Rev D 10:2445
77. Car R, Parrinello M Phys Rev Lett 55:2471
78. Rosensweig RE (1985) Ferrohydrodynamics. Cambridge University Press, Cambridge
79. Wang ZW, Holm C (2001) J Chem Phys 115:6277
80. Bródka A (2004) Chem Phys Lett 400:62
81. Toukmaji A, Sagui C, Board J, Darden T J Chem Phys 113:10913
82. Tcl/Tk homepage (2003) http://tcl.activestate.com/
83. LAM/MPI homepage (2004) http://www.lam-mpi.org/
84. MPICH homepage (2004) http://www-unix.mcs.anl.gov/mpi/mpich/
85. FFTW homepage (2003) http://www.fftw.org/
86. CVS concurrent versions system homepage (2003) http://www.cvshome.org/

Adv Polym Sci (2005) 185: 111–162
DOI 10.1007/b136795

Published online: 23 July 2005

Simulation of Charged Colloids in Solution

Per Linse

Physical Chemistry 1, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, S-221 00 Lund, Sweden
Per.Linse@fkem1.lu.se

Abstract Physicochemical properties of solutions of charged colloids are often dominated by the electrostatic interactions present in such systems. A full determination of these properties constitutes a highly nontrivial many-body problem involving long-range interactions. In electrostatically strongly coupled systems, it is essential to explicitly include both the charged colloids and the small ions in the model. The present review describes recent advances in performing Metropolis Monte Carlo simulations of such systems modeled within the primitive model of electrolytes. Four representative colloidal systems are systematically used in combination with three different boundary conditions. First, the spherical cell model is considered, and it is used primarily to examine the distribution of the counterions near a colloid. Second, the cylindrical cell model containing two colloids is presented, and it is employed to calculate the mean force and/or the potential of mean force acting on one of the colloids. Several methods are utilized and their merits are compared. Finally, full structural and thermodynamic properties are presented by using a cubic simulation box with periodic boundary conditions applied. An account of the system size convergence, the convergence of the Ewald summation, including an estimate of truncation errors and practical guidelines, and the ability to increase the simulation efficiency by using cluster trial displacements is provided.

Keywords Charged colloids · Monte Carlo simulation · Mean force · Ewald summation · Cluster trial move

Abbreviations

DLVO	Derjaguin–Landau–Verwey–Overbeek
MI	minimum image
MC	Monte Carlo
PB	Poisson–Boltzmann
SC	spherical cutoff
pmf	potential of mean force
RDF	radial distribution function

1 Introduction

Charged colloids in solution are ubiquitous in a wide variety of biological and technical systems. Some examples are proteins made by amino acids, micelles formed by charged surfactants or charged block copolymers, microemulsions formed by water, oil, and charged surfactants, silica particles made by silica oxide, and polystyrene based latex particles. In these systems, the physicochemical properties are to a large degree determined by electrostatic forces. Despite extensive studies of these forces for the last 50 years, the electrostatic interactions in such systems remain a central problem in colloidal science [1, 2].

The classical DLVO theory [3, 4] describes the interaction between two charged colloids in a solution, and this theory constitutes a cornerstone in colloidal science. The DLVO potential contains two terms, originating from

van der Waal and electrostatic interactions, respectively. The electrostatic one predicts a purely repulsive force between like-charged colloids and has often satisfactorily explained the predominant charge stabilization in these systems. However, the DLVO theory is now challenged by the fact that attractive interactions between like-charged colloids or like-charged planar surfaces in salt-free or nearly salt-free solutions have been found or indirectly inferred in a growing number of experimental studies [5–10]. The nature of the colloidal attraction appears to be of (at least) two kinds, one appearing only at short separations (< 10 Å) [5, 6] and the other only at longer separations (≈ 1 μm) [7–10].

The primitive model of electrolytes constitutes a firm basis for examining the intercolloidal structure of colloidal solutions as well as the distribution of small ions near charged colloids. In this model, the charged colloids (referred to as macroions) and the small ions are both represented as charged species, whereas the solvent is treated as a dielectric medium. From simulations that provide essentially the exact solution of such a model system, it has been found that the primitive model of electrolytes is able to predict an electrostatically driven attraction between like-charged colloids. If mathematical approximations are involved, the appearance of an attraction is still possible in, e.g., the hypernetted-chain approximation or is lost in mean-field theories like the PB approximation. Often, the description of colloidal solutions is simplified by making a preaverage over the small ions. Typically, theories based on such a description predict a repulsive force between like-charged colloids at all conditions. Hence, in addition to providing a detailed description of the structure of colloidal solutions, simulation results of the primitive model of electrolytes are also of great importance for assessing the accuracy of (i) approximate statistical-mechanical theories and (ii) simpler models. These aspects are further discussed in recent reviews [11–14].

Employing the primitive model of electrolytes, attractive forces between like-charged colloids arising from electrostatic interactions have been found in computer simulations [15–28] and proposed by various theories [29–37]. In most of the simulation studies reported and in some of the theoretical investigations, the attractive forces seem to be of a short-range character originating from correlation effects [15–22, 25–31, 37]. In some other simulation studies, the attraction is due to metastable counterion states [23, 24]. Other theoretical investigations predict a more long-range attraction [32–36], however not yet having been corroborated by simulations. In some of the theoretical approaches, attraction only appears when a simple salt is added.

Simulation studies and theoretical developments of other systems show that short-range attraction also may appear between like-charged planes [38–40], cylinders [37, 41–48], and stiff chains [49]. Attraction arising from electrostatic interactions among like-charged flexible objects has also been documented. For example, such an attraction gives rise to chain collapse [50–53] and phase instability [54] in systems containing flexible

polyelectrolytes, to the formation of toroids in systems containing semiflexible polyelectrolytes [55], and collapse of cross-linked polyelectrolyte gels [56, 57].

The aim of this review is to provide some state-of-the-art knowledge on the methodology of simulating equilibrium properties of solutions of charged colloids using the primitive model of electrolytes. The review is primarily based on experiences accumulated in this area during more than two decades. To make the review as coherent as possible, properties and simulation aspects, focused mainly on four representative systems, will be given. Consequently, most of the presented data are original.

Besides selecting a model, boundary conditions have to be applied, and the preferred one will depend upon the properties of interest. The most common ones are a spherical cell containing one macroion, a cylindrical cell containing two macroions, and a cubic box with periodic boundary conditions. The first is most frequently used to examine the distribution of small ions near a macroion and approximate thermodynamic results, the second to determine the force between two macroions mediated by the small ions, and the last boundary condition provides full structural and thermodynamic properties of a bulk system.

The most common simulation method in this area is Metropolis MC simulation, but molecular dynamics and Brownian dynamics are also employed. As mentioned above, simulations provide essentially the exact solution of the given model systems, but the results are subjected to statistical uncertainties. Nevertheless, the statistical uncertainties can normally be reduced to an arbitrary low level by increasing the simulation length. More severe are other methodological difficulties that accelerate as the size and charge asymmetries between the colloidal particles and the small ions are increased. These difficulties originate from (i) the long-range nature of the strong electrostatic interactions and (ii) the spatial accumulation of a large number of counterions near the macroions [58]. The so-called Ewald summation, based on ideas of Ewald proposed 80 years ago [59], has during the last two decades been an established method to handle long-range interactions. More recently, cluster move techniques have found to be very profitable for accelerating the convergence of the simulation of highly asymmetric electrolytes [60, 61]. Up to a 10^6-fold reduction of computer time can be achieved for typical model systems.

The organization of the review is as follows. In Sect. 2, the primitive model of electrolytes is introduced. A set of reduced parameters is discussed, and the connection to experimental systems is provided. Three different boundary conditions are presented. In Sect. 3, some general simulation aspects are given. The properties of four specific systems, representing different experimental cases, are provided in Sect. 4. Section 5 focuses on the calculation of the interaction between two macroions mediated by their counterions. Three numerical approaches will be examined and their relative merit compared. In

Sect. 6, different aspects on the simulation of bulk solutions of asymmetric electrolytes using periodic boundary conditions are given. First a comparison of different boundary conditions is presented in Sect. 6.1. Then, in Sect. 6.2, the Ewald summation is examined, and issues such as system size convergence, energy summation convergence, and optimization of the CPU time are discussed. Section 6.3 contains an analysis of the selection of trial moves, and in particular the usefulness of a cluster move technique is illustrated. Furthermore, a second-level cluster move technique, facilitating simulation of phase-separating systems, will also be treated briefly.

2 Model

2.1 The Primitive Model of Electrolytes

The primitive model of electrolytes constitutes a firm basis for statistical-mechanical description of solutions of charged colloids. This model will be adopted throughout, and it originates from the more general McMillan–Mayer solution theory [62, 63].

Within the primitive model of electrolytes, ionic species are represented by charged hard spheres differing in charge and size, whereas the solvent enters the model only through its relative permittivity. For simplicity, solutions containing only charged colloids and their counterions forming an electroneutral system are considered in this review, hence no simple salt is added. In what follows, colloids will be referred to as macroions. The interactions between particles are assumed to be pairwise additive, and for pair ij, where i and j denote either a macroion (M) or a counterion (I), it is given by

$$u_{ij}(r) = \begin{cases} \infty\,, & r < R_i + R_j\,, \\ \dfrac{Z_i Z_j e^2}{4\pi\varepsilon_0\varepsilon_r}\dfrac{1}{r}\,, & r \geq R_i + R_j\,, \end{cases} \tag{1}$$

where Z_i is the charge of particle i, R_i the radius of particle i, e the elementary charge, ε_0 the permittivity of vacuum, ε_r the relative permittivity of the solvent, and r the center-to-center separation between the particles. The model is schematically illustrated in Fig. 1.

The description of such a colloidal solution requires, in total, eight physical parameters, viz., Z_M, Z_I, R_M, R_I, ρ_M, ρ_I, T, and ε_r, where ρ_i is the number density of particles i and T the absolute temperature of the system. Not all of the parameters are independent. Here, (i) application of electroneutrality, (ii) noticing that $Z_i Z_j e^2/(4\pi\varepsilon_0\varepsilon_\mathrm{r} kT)$ appears collectively in the partition function, and (iii) expressing all lengths in, say, R_M, decreases the number of

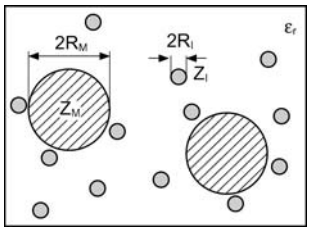

Fig. 1 Schematic illustration of the primitive model of electrolytes including spherical macroions with specified charge Z_M and radius R_M and spherical counterions with charge Z_I and radius R_I in a solution characterized by a relative permittivity ε_r

independent parameters by 1, 2, and 1, respectively, i.e., a decrease of in total four parameters. Hence the system is fully characterized by four reduced parameters. Moreover, since in colloidal systems $R_I \ll R_M$, the reasonable simplification $R_I \equiv 0$ is made, which results in only three remaining reduced parameters. The set consisting of (i) the macroion-charge-to-counterion-charge ratio Z_r, (ii) the macroion volume fraction ϕ_M, and (iii) an electrostatic coupling parameter Γ_{II} will be used, although of course other sets are possible. These parameters are defined as

$$Z_r \equiv -(Z_M/Z_I), \tag{2}$$

$$\phi_M \equiv (4\pi/3)R_M^3\rho_M, \tag{3}$$

$$\Gamma_{II} \equiv Z_I^2 l_B/R_M, \tag{4}$$

where $l_B \equiv e^2/(4\pi\varepsilon_0\varepsilon_r kT)$ is the Bjerrum length denoting the separation between two unit charges at which their electrostatic energy becomes equal to the thermal energy and with k being the Boltzmann constant.

Other length variables commonly of interest are the Debye screening length $\lambda_D \equiv [(eZ_I)^2\rho_I/(\varepsilon_0\varepsilon_r kT)]^{-1/2}$ with $\rho_I = Z_r\rho_M$, and where only the counterions are supposed to contribute to the screening, and the Wigner–Seitz cell radius $r_0 \equiv [(4\pi/3)\rho_M]^{-1/3}$. The former denotes the decay length of the screened Coulomb interaction, and the latter is a measure of the typical separation between two repelling and neighboring macroions. Expressed in the length unit R_M and reduced parameters, they become $\lambda_D/R_M = (3Z_r\Gamma_{II}\phi_M)^{-1/2}$ and $r_0/R_M = \phi_M^{-1/3}$, respectively. Moreover, we notice that the set (Z_r, Γ_{II}) can be combined into (Γ_{MI}, Γ_{II}) with $\Gamma_{MI} \equiv Z_r\Gamma_{II}$. With this new set the PB theory, where the counterion-counterion correlations are neglected, is obtained in the limit $\Gamma_{II} \to 0$ at nonzero Γ_{MI}. This is consistent with recent field-theoretical arguments [37], which say that the PB theory is recovered in the limit $\Xi \equiv Z_M Z_I^2 l_B^2/(2R_M^2) = \Gamma_{MI}\Gamma_{II}/2 \to 0$ at finite reduced Gouy–Chapman length $\mu/R_M \equiv 2R_M/(Z_M Z_I l_B) = 2/\Gamma_{MI}$.

The case of an aqueous solution at room temperature containing monovalent counterions is commonly encountered, and for convenience this case

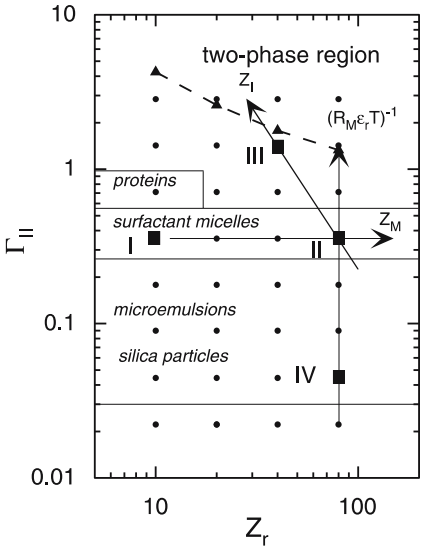

Fig. 2 Overview of the (Z_r, Γ_{II})-parameter plane containing the location of typical charged colloids with monovalent counterions in aqueous solution at room temperature and the location of Systems I–IV (*squares*). The directions of increased value in (i) counterion charge Z_I, (ii) macroion charge Z_M, and (iii) the combined parameter $(R_M\varepsilon_r T)^{-1}$ are shown by *arrows*. Also shown are other simulated systems (*dots*) and the gas–liquid coexistence curve (*dashed curve with triangles*) [22]

will be referred to as the standard condition. Figure 2 shows the approximate location in the (Z_r, Γ_{II}) plane at the standard condition of different important experimental systems including (i) solutions of charged and globular proteins, (ii) solutions of charged micelles formed by surfactants, (iii) microemulsion systems, and (iv) solutions of silica particles. The Z_r-range relevant for aqueous solutions of latex particles is, however, not covered. Although their lower size limit of ≈ 1000 Å, corresponding to $\Gamma_{II} \approx 0.01$ at the standard condition, is tractable, their lower bare charge $Z_M \approx 1000$ is still too large for systematic simulation studies.

In this review, properties and simulation aspects of mainly four colloidal systems will be considered. The location of these systems in the (Z_r, Γ_{II}) plane are given by filled squares in Fig. 2, and their full characterization in terms of (Z_r, ϕ_M, Γ_{II}) is provided in Table 1. For an easier comparison among the systems, the macroion volume fraction $\phi_M = 0.01$ is used throughout. Obviously *each* of the four systems represents several experimental ones. For example, Systems I–III could represent an aqueous solution of charged surfactant micelles with radius $R_M = 20$ Å, where the micelles in System I could be formed by a mixture of ionic and nonionic surfactants with monovalent counterions, in System II by ionic surfactants with monovalent counterions, and in System III by ionic surfactants with divalent counterions. Finally, the

Table 1 Definition of the four systems considered in terms of the reduced units Z_r, Γ_{II}, and ϕ_M, typical experimental colloidal systems that they could represent, and two other reduced parameters

System	Z_r	Γ_{II}	ϕ_M	Typical experimental systems[a] description	Z_M	Z_I	R_M Å	Γ_{MI}^b	Ξ^c
I	10	0.3558	0.01	mixed surfactant micelle with monovalent counterions	– 10	+ 1	20	3.558	0.633
II	80	0.3558	0.01	surfactant micelle with monovalent counterions	– 80	+ 1	20	28.46	5.06
III	40	1.423	0.01	surfactant micelle with divalent counterions	– 80	+ 2	20	56.92	40.5
IV	80	0.0445	0.01	silica particle	– 80	+ 1	160	3.558	0.0792

[a] Aqueous solution ($\varepsilon_r = 78.3$) and room temperature ($T = 300$ K) have been assumed.
[b] Macroion–counterion potential energy at contact, $\Gamma_{MI} \equiv Z_r \Gamma_{II}$.
[c] Field-theoretical coupling parameter $\Xi = \Gamma_{MI}\Gamma_{II}/2$ (in the present notation) [37].

macroion charge in System IV could be the same as in Systems II and III, but the electrostatic coupling is weaker by, e.g., making the macroion radius larger, and then this system could represent a suspension of particles with a lower surface charge density as compared to System II. These examples of experimental realizations are also summarized in Table 1.

2.2 Boundary Conditions

Besides the description of the ionic species and their interactions, a boundary confining the particles needs to be provided. The choice of such a boundary condition is dictated by the properties of interest to be investigated. Here, (i) a spherical cell, (ii) a cylindrical cell, and (iii) a cubic box with periodic boundary conditions will be considered. Figure 3 illustrates these boundary conditions and also displays typical position of the particles.

The *spherical cell* approach is based on a hypothetical subdivision of the colloidal solution into electroneutral subvolumes, each containing one macroion and corresponding amount of counterions and solvent. This boundary condition is most frequently used to examine the distribution of small ions near a macroion and to obtain approximate thermodynamic results. It is most often implemented by placing a single macroion concen-

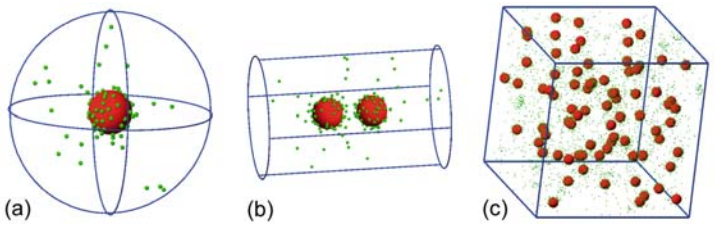

Fig. 3 Snapshots illustrating boundary conditions and typical particle positions of System II. **a** Spherical cell containing one macroion and 80 counterions. **b** Cylindrical cell containing two macroions and 160 counterions. **c** Cubic box with periodic boundary conditions containing 80 macroions and 6400 counterions. A radius $0.1R_M$ has here been assigned to the point counterions. In **a** $R_{sph} = 4.64R_M$, in **b** $R_{cyl} = 4R_M$ and $L_{cyl} = 12R_M$, and in **c** $L = 32.22R_M$, resulting in macroion volume fractions $\phi_M = 0.010$, 0.014, and 0.010, respectively

trically in the spherical cell and confining the small ions by the external potential

$$u(r) = \begin{cases} 0, & r \le R_{sph}\,, \\ \infty\,, & r > R_{sph}\,, \end{cases} \tag{5}$$

where $\boldsymbol{r} \equiv (x, y, z)$ is the position of particle i and R_{sph} the radius of the spherical cell. Hence, the small ions explore the volume of the sphere (Fig. 3a). Of course, the influence of the remaining solution is taken into account only in an averaged way. The spherical boundary condition is most useful when the macroions repel each other and becomes inappropriate when they possess attractive interactions.

A cell approach can also be used to determine the force between two macroions mediated by the small ions. For that purpose, a *cylindrical cell* will be employed; however, the exact shape of the cell is not important. In the cylindrical cell approach, two macroions are placed on the C_∞ axis of the cylinder, and the confining external potential, applying to all particles, is given by

$$u(r) = \begin{cases} 0, & \sqrt{(x^2 + y^2)} \le R_{cyl} \quad \text{and} \quad |z| \le L_{cyl}/2\,, \\ \infty, & \text{otherwise}\,, \end{cases} \tag{6}$$

where it has been assumed that the C_∞ axis is placed along the z-direction and where R_{cyl} denotes the cylinder radius and L_{cyl} the cylinder length (Fig. 3b). The macroion separation is either held fixed or variable, depending on the method used to calculate the mean interactions.

A major obstacle of simulating a bulk liquid confined in a container is the large fraction of particles that are in contact with the walls of the container. This surface effect can be circumvented by applying *periodic boundary conditions* in all three directions. For example, a periodic boundary condition in

one dimension, say, in the x-direction, involves the replacement of the separation x ($x > 0$) between two particles with $L - x$, where L is the box length, if x exceeds $L/2$ [64]. Thus, the effect of confining surfaces is eliminated and all spatial positions in the box become identical. However, still some system size dependence remains (see discussion in Sect. 6.2.2). Occasionally, this boundary condition is referred to as toroidal boundary conditions. Figure 3c illustrates a cubic box containing macroions and counterions to which periodic boundary conditions have been applied.

3 General Simulation Aspects

The Metropolis MC method in the canonical ensemble is the most frequently used simulation approach to solve the primitive model of electrolytes. Averages of static properties are taken from a large set of Boltzmann-weighted configurations. Molecular dynamics and Brownian dynamics constitute two other methods to determine static and dynamic properties of molecular systems. Their implementations are, however, complicated for systems possessing impulsive forces in combination with other forces. Hence, a soft-sphere repulsion is frequently used instead of the hard-sphere one when simulating such systems with these methods.

In the spherical and cylindrical cell approaches, all pair interactions according to Eq. 1 are taken into account and the execution of such simulations is often straightforward. Simulations of fluids confined in a box with periodic boundary conditions involve a larger number of particles with a concomitant increase of computer load, typically 50 to 100 macroions, their counterions, and salt species (if any) are used. As already alluded to in the introduction, (i) the treatment of the long-range electrostatic interactions and (ii) the ability to generate Boltzmann-weighted configurations spanning the relevant part of the configurational space are two important issues in such simulations. These issues will be elaborated in Sect. 6. Regarding other simulation-related issues, including determination of equilibrium and of precisions of mean values from production simulations through block averages and autocorrelation functions, the reader is referred to [64–66].

Unless otherwise stated, the properties of Systems I–IV are based on equilibrated systems and made by MC simulations. Uncertainties given are based on block averages. The length of the production runs are given in number of MC passes, N_{pass}; such a pass comprises, on average, one trial move per particle.

4
Properties of the Model

This section contains a brief resumé of the use of the three different boundary conditions, selected equilibrium properties of asymmetric electrolytes based on Systems I–IV, and other recent advances.

4.1
Counterion Accumulation

The spherical cell model was used early to examine the distribution of the small ions near a charged macroion and the thermodynamics of such systems. In the past, the model was used to investigate, e.g., the electrostatic potential and the small ion distribution [67–74], the self-diffusion of counterions [75–78], the micelle formation of charged surfactants [79], and full phase diagrams of charged surfactant-water systems [80]. Many of those investigations were based on the PB equation, providing an approximate description of the model system.

The radial counterion distribution, expressing the local density of the small ions as a function of the distance from the macroion, is the principal result of the spherical cell model. Figure 4 displays such distributions for Systems I–IV, and counterion concentrations at the macroion and cell surfaces are collected in Table 2.

Generally, there is an accumulation of the counterions near the macroion surface with the maximal concentration appearing at contact, $r = R_M$ (remember $R_I = 0$), arising from the attractive macroion–counterion electrostatic interaction. In Systems I and IV, the counterion distribution deviates only moderately from a uniform distribution, which originates from their comparable small Γ_{MI} values (Table 1). The distribution of the counterions becomes much more inhomogeneous in Systems II and III, representing solutions of charged surfactants with monovalent and divalent counterions, respectively. For System III, the local concentration at the macroion surface is nearly 1000 times higher than the average one (Table 2). Returning to Systems I and IV, their counterion distributions are indistinguishable within the statistical uncertainty. This is a consequence of the fact that (i) the values of Γ_{MI} are identical for these systems and (ii) the values of Γ_{II} are small (< 1) (Table 1 and the discussion in Sect. 2.1).

A comparison between results from the MC simulations and the PB theory shows that they are indistinguishable for Systems I and IV, slightly different for System II, and considerably different for System III. This increasing difference is consistent with an increasing value of Ξ for System IV, System I, System II, and System III (Table 1). In particular, the PB approximation underestimates the accumulation of the small ions near the macroion

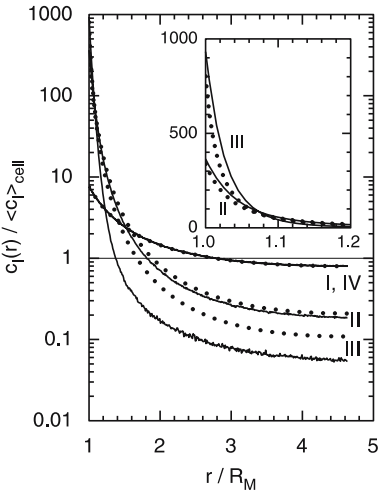

Fig. 4 Reduced counterion concentration profile for Systems I–IV as a function of the radial distance evaluated from the spherical cell model containing one macroion and Z_r counterions solved by MC simulations (*solid curves*) and the PB equation (*dotted curves*). The data for Systems I and IV practically overlap (MC simulation) or are identical (PB equation). The *thin solid line* corresponds to a uniform distribution. The *insert* shows an expansion for short distances for Systems II and III. $R_{cell} = 4.64R_M$, $N_{pass} = 10^6$, and a histogram width of $0.01R_M$ for sampling the simulated profiles

Table 2 Reduced counterion concentration at the macroion and cell surfaces evaluated from the cell model using MC simulation and the PB equation[a]

System	$c_I(R_M) / \langle c_I \rangle_{cell}$ MC[b]	PB[c]	$c_I(R_{cell}) / \langle c_I \rangle_{cell}$ MC[b]	PB[c]
I	7.8(2)	7.55	0.80(3)	0.799
II	364(3)	337	0.186(3)	0.209
III	932(5)	802	0.054(2)	0.109
IV	7.6(1)	7.55	0.801(6)	0.799

[a] $\langle c_I \rangle_{cell}$ denotes the average counterion concentration in the cell. $R_{cell} = 4.64R_M$ and $N_{pass} = 10^6$.
[b] Values obtained by using a histogram width $0.01R_M$ and an extrapolation using a harmonic fit.
[c] In the PB approximation, Systems I and IV become identical when described in reduced units.

(see insert of Fig. 4), leading to a too large concentration further away. Furthermore, within the spherical cell model, the osmotic pressure is given by $P = kTc(R_{cell})$, where $c(R_{cell})$ is the total concentration of the small ions at the cell boundary [81]. This relation holds (nontrivially) also for the PB approximation [82]. The discrepancies of the cell contact values between the PB

approximation and the exact ones given in Table 2 show that the PB equation overestimates the osmotic pressure by 10% for System II and 100% for System III. This confirms the established view that the PB equation provides an acceptable description of aqueous solution at ambient temperature containing charged surfaces with monovalent counterions, but *not* with divalent ones.

More recently, the spherical cell model has been used to examine the effects of replacing the homogeneous surface charge density with discrete macroion surface charges [83–85] and with more detailed protein features [86]. Another use of the cell model is to assess the importance of surface polarization appearing at the macroion–solvent interface when taking into account the lower permittivity inside the macroion [83, 87, 88]. Finally, an eccentric positioning of the macroion in a spherical cell model was used to investigate effective macroion–macroion interactions [89].

The complexation between one macroion and an oppositely charged polymer has also been addressed by MC simulations using the cell model. Early simulations were focused on the thermodynamics of the complexation under different conditions [90–92], whereas more recent simulations have complemented theoretical predictions [93, 94] and mapped the diversity of the macroion–polyion complex structure in more detail [95, 96].

4.2 Mean Force

Other important aspects of asymmetric electrolytes can be obtained from the mean force acting between two macroions mediated by the small ions. Such determinations have been performed for structureless macroions for equilibrated counterions [15–17, 19, 25, 26, 28, 97] and metastable counterion states [23, 24]. In these investigations, different boundary conditions have been applied, from cell models to cubic boxes with periodic boundary conditions. A similar study for two macroions with polymeric counterions has also been performed [98]. Since the macroions and their counterions are confined to a finite volume, which approximately takes into account the influence of the surrounding electrolyte, the force obtained is a mean force rather than an effective one [13].

The mean force and the potential of mean force acting on one of the two macroions as a function of the macroion separation r for Systems I–IV obtained from the cylindrical cell model are provided in Fig. 5. Starting with System II, representing a solution of surfactant micelles with monovalent counterions, the mean force is purely repulsive at all separations (Fig. 5a). The eightfold reduction of the macroion charge (System II → System I) leads to a strong reduction of the magnitude of the repulsive mean force. The replacement of monovalent counterions with divalent ones (System II → System III) also strongly reduces the long-range repulsion, and, moreover, at

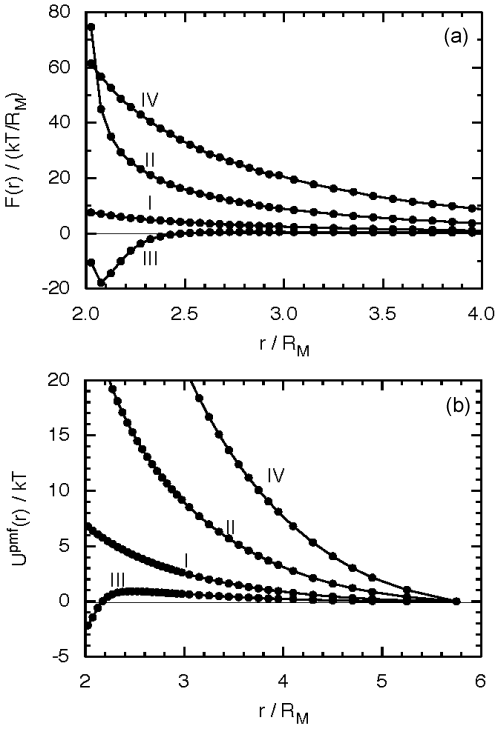

Fig. 5 **a** Mean force and **b** potential of mean force for Systems I–IV as a function of the macroion separation evaluated from the cylindrical cell model containing two macroions and $2Z_r$ counterions. In **b**, $U^{pmf}(r = 5.75R_M) = 0$ has been adopted. $R_{cyl} = 4R_M$, $L_{cyl} = 12R_M$, and $N_{pass} = 10^6$

short separation an *attractive* mean force appears. Finally, a reduction of the electrostatic coupling through an increase of macroion radius, permittivity, or temperature (System II → System IV) gives rise to an *increase* of the long-range repulsion.

The corresponding potential of mean force (pmf) displays a similar functional form (Fig. 5b). In Systems II and IV, the mean force has not approached zero at the longest separation considered, and hence U^{pmf} should be considerably more repulsive (shifted upward) than shown to match the conventional limit $U^{pmf}(r) \to 0$ as $r \to \infty$. For System I, the pmf at contact ($r = 2R_M$) is reduced to $\approx 7kT$. Regarding System III, only a weak long-range repulsion remains, a barrier of $\approx 1kT$ appears at $r \approx 2.4R_M$, and a global minimum of $\approx -2kT$ occurs at macroion contact.

Lately, mean forces and pmfs have been simulated for models containing discrete surface charges inhomogeneously distributed with a spherical hard-sphere potential [99–101] and protein models with a more irregular hard-core potential [102].

4.3 Structure

The first published simulation of solution properties of substantial size and charge asymmetric electrolytes was made over 20 years ago and was restricted to low charge asymmetry, $Z_r = 12$ [103]. Despite much improved computer resources, later simulation studies were still bounded by an upper charge asymmetry limit of $Z_r \approx 20$ [21, 58, 74, 104–108]. It was first with the use of an improved sampling technique that solutions of asymmetric electrolytes with larger charge asymmetries could efficiently be treated [18, 20, 22, 27, 109].

Representative radial distribution functions (RDFs), $g_{ij}(r)$, again for Systems I–IV, are shown in Fig. 6. These functions display the relative density of a particle of type j at a distance r from a particle of type i. In homogeneous fluids, $g_{ij}(r)$ is conventionally defined to approach unity at large distances. At distances $r < R_i + R_j$, where hard-sphere overlap appears, $g_{ij}(r)$ is exactly zero.

Starting with System II, the macroion–macroion RDF displayed in Fig. 6a confirms the conventional view that macroions are well separated and form a fluid with an order extending at most a few macroions away. The closest macroion–macroion separation exceeds the contact separation $2R_M$ and the first maximum of the macroion–macroion RDF appears at $\approx \rho_M^{-1/3}$, demonstrating a strongly repulsive effective force acting between the macroions and consistent with the pmf given in Fig. 5b. The main effect of varying the macroion density is a shift of the maxima and minima maintaining the $\rho_M^{-1/3}$ relation [22]. A strong accumulation of the counterions near the macroions is clearly demonstrated in Fig. 6b, as was also deduced from the spherical cell model. The contact value of g_{MI} increases as the density decreases, showing a more uneven counterion distribution at lower densities [22]. However, the counterion number density at contact $\rho_I g_{MI}(r = R_M)$ is only weakly (roughly logarithmically) density dependent, an important factor for simulation optimization. Macroion–counterion RDFs obtained from simulations of fluids and counterion radial distribution from simulations using the spherical cell model are nearly identical at short separations where the influence from other macroions and their nearby counterions is small. Finally, the counterion–counterion RDF shown in Fig. 6c displays a positive density correlation in the range $0.1 - 2R_M$. The shape of $g_{II}(r)$ is a characteristic feature of dilute asymmetric electrolytes and reflects the combined effect of the repulsion among the macroions, the attraction between the counterions and the macroions, and the repulsion among the counterions themselves.

When the macroion charge is reduced (System II $\rightarrow$ System I), $g_{MM}(r)$ displays a reduced ordering of the macroions. Obviously, the effective repulsion among the weaker charged macroions is smaller, again consistent with the pmf from the cylindrical cell approach. The contact value of g_{MI} is re-

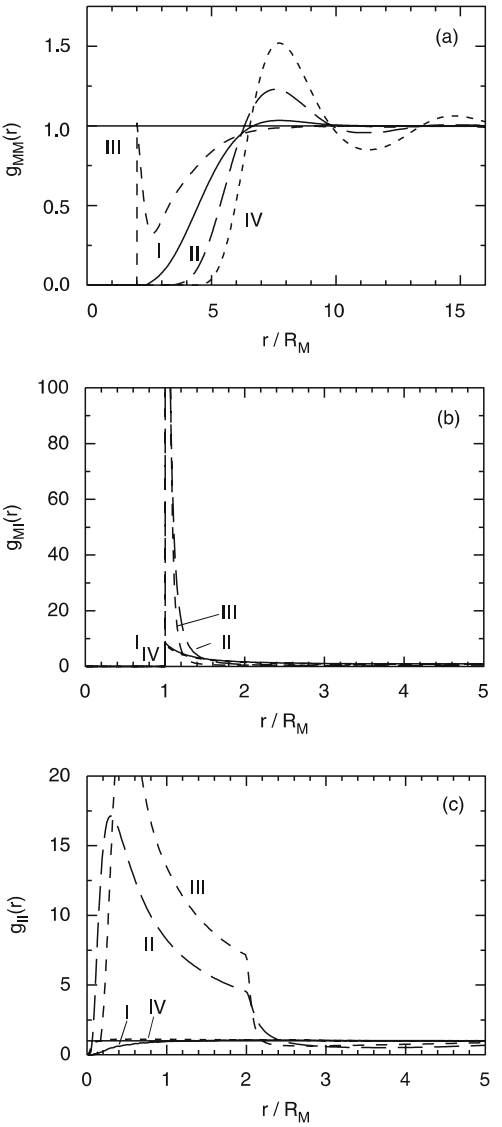

Fig. 6 **a** Macroion-macroion, **b** macroion-counterion, and **c** counterion–counterion radial distribution functions for Systems I–IV obtained from simulations using a cubic box containing 80 macroions and $80Z_r$ counterions with Ewald summation (Sect. 6.2) and macroion cluster moves (Sect. 6.3). *Thin solid lines* correspond to uniform distributions. $L = 32.22R_M$ and $N_{pass} = 5 \times 10^4$. Data taken from ref [22]

duced to ≈ 10, demonstrating a much weaker accumulation of counterions near the macroion, also an important factor for efficient MC sampling. The counterion–counterion RDF is very smooth, and no spatial correlation among

the counterions remains except for a repulsion at short distances. Hence, the reduced macroion charge leads to weakened structure among all pairs of particle types.

As the monovalent counterions are replaced by divalent ones (System II → System III), the shape of the macroion–macroion RDF is drastically changed. The characteristic maximum at $\approx \rho_M^{-1/3}$ has vanished, $g_{MM}(r) < 1$ for all separations $r < 10R_M$, and a sharp local maximum appears at $r \approx 2R_M$. Hence, the long-range repulsion is absent, and there is a tendency to form macroion pairs. Again the RDF seems to be consistent with the corresponding pmf results. In particular, a positive pmf at short separations (Fig. 5b) corresponds to a value of the macroion–macroion RDF of less than one (Fig. 6a) and the negative pmf at contact for System III with a maximum in the RDF. [In fact, defining $U^{pmf}_{fluid}(r)/kT \equiv -\ln[g_{MM}(r)]$, a close resemblance between $U^{pmf}_{fluid}(r)$ and $U^{pmf}(r)$ is obtained at conditions where the repulsion is not too large [26].] The macroion–counterion RDF shows that the accumulation of the counterions near the macroion is stronger, and the $g_{II}(r)$ displays a stronger spatial correlation among the counterions as compared to System II, as was also concluded from Fig. 4.

Finally, the changes in structure at an increase of the macroion radius from 20 to 160 Å (System II → System IV) will be considered. First, $g_{MM}(r)$ shows that the macroion structure is *enhanced*, again consistent with the pmf data. Since we consider systems at same volume fractions, this enhancement must have an electrostatic origin.

To summarize the observations, in the sequence (System IV → System II → System III) we observe (A) a decrease of the repulsive mean force between the two macroions and a reduced long-range structure (B) followed by the appearance of an effective attraction at short separation.

(A) The physical reason for the weakening of the macroion structure as, e.g., the macroion radius is reduced (System IV → System II), is a nonlinear response. Table 1 shows the macroion–counterion potential energy at contact amounts to $\Gamma_{MI} = 3.6$ and 28.5, respectively, and Fig. 6 shows a superlinear increase of the counterion accumulation. Although the direct macroion–macroion repulsion also increases as Γ_{II} is increased, the effect of the screening grows faster. In fact, the macroion structure *decreases* as the electrostatic coupling Γ_{II} is raised when $\Gamma_{II} > 5/Z_r$ [22].

(B) Moreover, as the electrostatic coupling parameter Γ_{II} is increased even further, here by replacing monovalent counterions with divalent ones, a situation occurs where the extension of the conventional double layer is very small, and a short-range attraction seems to appear. The appearance of such an attractive force at short separation is at odds with the DLVO theory [3, 4]. The origin of this attraction was discussed in some detail in [26]. Briefly, starting from a mean-field theory as the PB equation, where the spatial correlations among the counterions are neglected, only a repulsive interaction

is possible [110, 111]. With this perspective, the attraction arises from the spatial correlations among small ions residing on different macroions. The correlation effects come effectively into play only at such large Γ_{II}, where the conventional electrical double layer becomes very thin (essentially all counterions are electrostatically attracted to macroion surfaces).

At an even slightly larger electrostatic coupling than in System III, the solution becomes unstable. Figure 2 (dashed curve) displays the boundary of the unstable region in the (Z_r, Γ_{II}) plane at $\phi_M = 0.01$. Above the boundary the system separates into two phases, one rich and one poor in the electrolyte [22].

More complex systems such as solutions containing macroions and short flexible counterions have recently been simulated using the primitive model of electrolytes [112]. Solutions of macroions with simple counterions at different amounts of oppositely charged polyelectrolyte have also been investigated, and the sequence "complexation → phase separation → redissolution" was observed [113]. Similar simulations where the macroion represented lysozyme have also been performed [114]. Finally, by using a related soft-sphere model, the dynamics and, in particular, the self-diffusion of the macroions and the counterions have been investigated by employing Brownian dynamics simulation [115].

5 Cylindrical Cell Model

A brief theoretical background of the mean force and pmf evaluated in a cylindrical geometry will be given. Such an approach rests on the assumption that three and higher-order macroion correlations are of less importance. The performance of three different methods and a critical examination of the assignment of some selected parameters of these methods will be provided. Finally, some practical guidelines will be collected. Although the numerical examples are provided for spherical particles with hard-sphere and electrostatic interactions, mean forces and pmf's can easily be calculated for nonspherical particles and with particles possessing other types of interactions.

Mean forces and/or pmf's between macroions mediated by small ions, representing the situation in solution, have been calculated using a variety of boundary conditions: a cubic box with periodic boundary conditions [16, 17, 19, 100, 101], a cubic box with hard walls [15], one spherical cell [102], two spherical cells [97], and a cylindrical cell [26, 28]. The assumption of small higher-order correlation effects implies that the mean force should not strongly depend on (i) the boundary conditions or (ii) the shape of the cell or the box. The latter issue has briefly been examined for the cylindrical cell; otherwise these aspects have not yet been analyzed. Finally, the

cylindrical cell model possesses two favorable properties: (i) fairly equal distribution of the solution volume around the macroions and (ii) existence of several methods to evaluate the mean force or the pmf.

5.1 Mean Force and Potential of Mean Force

Consider two macroions and their counterions confined in a cylindrical cell as illustrated by Fig. 3b and with the cylinder symmetry axis in the z-direction. From general principles of local equilibrium, the mean force $\boldsymbol{F}$ acting on one of the macroions, say M, can be expressed as [13, 116, 117]

$$\boldsymbol{F} = \oiint_{\text{area}} \left[-kT\rho_{\text{I}}\boldsymbol{I} + \varepsilon_0\varepsilon_{\text{r}} \left(\boldsymbol{EE} - \frac{1}{2}\boldsymbol{E}^2\boldsymbol{I} \right) \right] \cdot \mathrm{d}\boldsymbol{s}\,, \tag{7}$$

where the integrand constitutes the stress tensor and the integration is over a closed (but otherwise arbitrary) area enclosing macroion M but excluding the other macroion. In Eq. 7, the first term denotes the kinetic contribution with ρ_{I} here being the small ion distribution and $\boldsymbol{I}$ the unit tensor, the second term denotes the contribution from the electrostatic interaction with $\boldsymbol{E}$ being the electrostatic field, and $\mathrm{d}\boldsymbol{s}$ denotes a surface element times a normal vector pointing away from macroion M. By symmetry, in the present context x and y components of the mean force are zero, and hence the mean force F lies in the z-direction. A positive value of F implies a repulsive and a negative value an attractive force. In what follows, r will denote the separation between the two macroions.

The related pmf U^{pmf} is defined by

$$U^{\text{pmf}}(r) \equiv -\int_{\infty}^{r} F(r')dr' \,. \tag{8}$$

At large r, $F(r)$ and $U^{\text{pmf}}(r)$ approach zero. Both quantities depend on (i) the direct repulsive macroion–macroion interaction $u_{\text{MM}}(r)$ and (ii) an indirect contribution mediated by the counterions.

5.2 Evaluation of Mean Force

As a consequence of the freedom of selecting the integration surface in Eq. 7, the mean force can be evaluated in an infinite number of ways. Two of them will be considered here.

5.2.1 Surface Approach

If the hard-sphere surface of macroion M is selected as the integration area, Eq. 7 can be expressed as

$$F(r) = F_{\mathrm{hs}}(r) + F_{\mathrm{elec}}(r)\,, \tag{9}$$

where

$$F_{\mathrm{hs}}(r) = \hat{\boldsymbol{z}} \cdot \left[-kT \oint\limits_{\substack{\text{macroion} \\ \text{surface}}} \rho_{\mathrm{MI}}(A)\, \mathrm{d}\boldsymbol{s} \right], \tag{10}$$

$$F_{\mathrm{elec}}(r) = \sum_{i \neq \mathrm{M}}^{N} \left\langle -\nabla_{r_{\mathrm{M}i}} u_{\mathrm{M}i}^{\mathrm{elec}}(r_{\mathrm{M}i}) \right\rangle . \tag{11}$$

In Eq. 10, $F_{\mathrm{hs}}(r)$ represents the net transfer of linear moments in the z-direction from small ions colliding with macroion M with $\rho_{\mathrm{MI}}(A)$ denoting the contact density of the small ions at the macroion surface. Since $\rho_{\mathrm{MI}}(A)$ varies over the macroion surface, $F_{\mathrm{hs}}(r)$ is not trivially zero. In Eq. 11, $F_{\mathrm{elec}}(r)$ represents the electrostatic force in the z-direction acting on macroion M arising from the electrostatic interaction with the other particles in the system.

The evaluation of the surface integral in Eq. 10 is numerically cumbersome, since the contact density as a function of the angle between $\hat{\boldsymbol{z}}$ and the surface normal has to be sampled. However, that can be circumvented by using an elegant method proposed by Wu et al. [17, 19]. Basically, $F_{\mathrm{hs}}(r)$ can be reexpressed as

$$F_{\mathrm{hs}}(r) = -kT \left[\lim_{\Delta r \to 0^+} \frac{P_{\mathrm{overlap}}(\Delta r)}{\Delta r} - \lim_{\Delta r \to 0^-} \frac{P_{\mathrm{overlap}}(\Delta r)}{\Delta r} \right], \tag{12}$$

where $P_{\mathrm{overlap}}(\Delta r)$ denotes the probability of at least one hard-sphere overlap appearing between the macroion and one small ion when the macroion is displaced the distance Δr in the z-direction. The two terms in Eq. 12 denote the transfer of linear moments in the negative and positive z-direction, respectively. According to Wu et al. [19], $P_{\mathrm{overlap}}(\Delta r)$ could be approximated with the average number of hard-sphere overlaps for sufficiently small Δr. Since such an approximation does not possess any numerical advantage, it has not been made in the examples given below.

The limits in Eq. 12 have to be evaluated numerically using a finite Δr. The determination of a reasonable Δr requires some numerical tests. A small Δr gives rise to a principally improved F_{hs}, but suffering from large (statistical) uncertainties, whereas a large Δr leads to a result with high accuracy, but possessing a systematic error. The largest Δr, at which F_{hs} still does not suffer

significantly from a systematic error (referred to as the optimal Δr), depends on (i) the system, (ii) macroion separation, and (iii) simulation length. Some of these aspects are illustrated in Fig. 7a, which displays F_{hs} with uncertainty bars as a function of Δr calculated for System II at macroion separation $r = 2.025R_{\mathrm{M}}$. The figure confirms the appearance of a systematic underestimation of F_{hs} at large Δr, and the increase in the uncertainty as Δr is reduced. The optimal Δr appears at ≈ 0.0002. Table 3 provides the optimal Δr and corresponding uncertainty $\sigma(F_{\mathrm{hs}})$ for Systems I–IV at $r = 2.025R_{\mathrm{M}}$ and $2.975R_{\mathrm{M}}$. We conclude that (i) the optimal Δr increases as r is increased, varies between 0.0002 and 0.005 for the shorter macroion separation, and becomes larger for the longer macroion separation, and (ii) the optimal Δr is smallest for System II. Wu et al. suggested $\Delta r = 0.02R_{\mathrm{M}}$ as the initial value for the search of the optimal Δr [19].

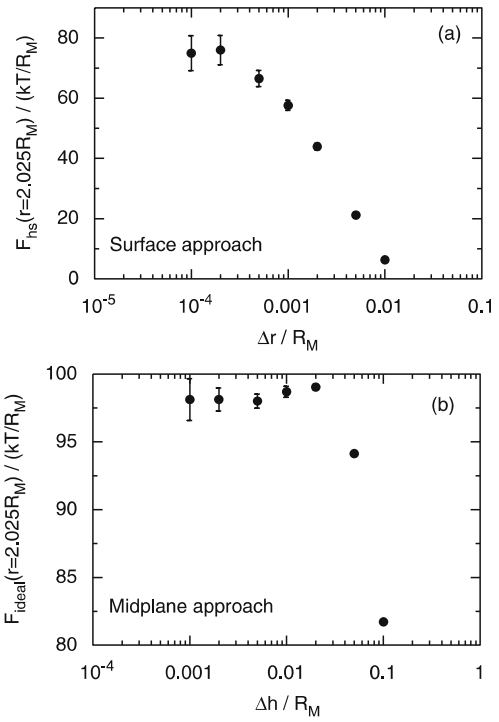

Fig. 7 **a** Hard-sphere contribution to the mean force F_{hs} as a function of Δr and **b** the ideal contribution to the mean force F_{ideal} as a function of Δh with uncertainty bars (one standard deviation) for System II at macroion separation $r = 2.025R_{\mathrm{M}} \cdot R_{\mathrm{cyl}} = 4R_{\mathrm{M}}, L_{\mathrm{cyl}} = 12R_{\mathrm{M}}$, and $N_{\mathrm{pass}} = 10^6$. The uncertainty estimates are based on a division of the simulations into 20 blocks. See Eqs. 12 and 16 for the meaning of Δr and Δh, respectively

Table 3 Mean-force calculations: optimal Δr and corresponding $\sigma(F_{hs})$ for the surface approach and optimal Δh and corresponding $\sigma(F_{ideal})$ for the midplane approach in units of R_M and kT/R_M at two different macroion separations. [a] See text for further details

	Surface approach				Midplane approach			
	$r = 2.025R_M$		$r = 2.975R_M$		$r = 2.025R_M$		$r = 2.975R_M$	
System	Δr	$\sigma(F_{hs})$	Δr	$\sigma(F_{hs})$	Δh	$\sigma(F_{ideal})$	Δh	$\sigma(F_{ideal})$
I	0.005	0.05	0.01	0.03	0.05	0.04	0.1	0.03
II	0.0002	5	0.002	1	0.02	0.2	0.1	0.03
III	0.002	2	0.01	0.1	0.02	0.2	0.1	0.01
IV	0.002	0.3	0.01	0.07	0.02	0.2	0.1	0.06

[a] $R_{cyl} = 4R_M$, $L_{cyl} = 12R_M$, and $N_{pass} = 10^6$

5.2.2 Midplane Approach

The surface enclosing one half of the cylinder constitutes another reasonable choice of the integration area to be used in Eq. 7. With this choice, the mean force can be divided into terms according to

$$F(r) = F_{ideal}(r) + F_{elec}(r)\,, \tag{13}$$

where

$$F_{ideal}(r) = kT\left[\rho_I(z=0) - \rho_I(z = L_{cyl}/2)\right]A\,, \tag{14}$$

$$F_{elec}(r) = \sum_{i<j}^{N}{}' \left\langle -\nabla_{r_{ij}} u_{ij}^{elec}(r_{ij}) \right\rangle\,. \tag{15}$$

In Eq. 14, $F_{ideal}(r)$ arises from the difference in the transfer of linear moments of the small ions across the planes $z = 0$ and $z = L_{cyl}/2$ with $\rho_I(z = z')$ being the small ion number density in the plane $z = z'$ averaged over the cross-section area A of the cylinder. As long as $r \ll L_{cyl}/2$, $F_{ideal}(r)$ is dominated by the $\rho_I(z = 0)$ term. In Eq. 15, $F_{elec}(r)$ denotes the average force operating across the plane $z = 0$ originating from the electrostatic interaction among the charged particles and the prime that the summation only includes pairs of particles located on different sides of the plane $z = 0$. Note, $F_{elec}(r)$ represents different quantities in the surface and midplane approaches.

Also, the evaluation of the densities at the z-planes in Eq. 14 requires a numerical protocol. A simple approach is to use

$$F_{ideal}(r) = kT \lim_{\Delta h \to 0^+} \left[\frac{\langle N_I \rangle_{\Delta h}^{(0)}}{\Delta h} - \frac{\langle N_I \rangle_{\Delta h/2}^{(-)} + \langle N_I \rangle_{\Delta h/2}^{(+)}}{\Delta h} \right]\,, \tag{16}$$

where $\langle N_I \rangle_{\Delta h}^{(0)}$ denotes the average number of small ions in a slab with the width Δh centered at $z = 0$ and $\langle N_I \rangle_{\Delta h/2}^{(\mp)}$ the average number of small ions in a slab with the width $\Delta h/2$ centered at $z = \mp(L_{cyl} - \Delta h/2)/2$. As with Δr in the surface approach, a small Δh provides a principally improved $F_{ideal}(r)$ but subjected to a large statistical uncertainty, whereas a larger Δh reduces the statistical uncertainty but increases the systematic error. Figure 7b shows F_{ideal} with uncertainty bars as a function of Δh calculated for System II at macroion separation $r = 2.025R_M$, and indeed such a behavior is observed. The optimal Δh is ≈ 0.02. Table 3 also contains optimal Δh and corresponding $\sigma(F_{ideal})$ values for Systems I–IV at two macroion separations. It is found that the optimal Δh (i) increases as r is increased and (ii) is less dependent than Δr on the particular system.

5.3
Evaluation of Potential of Mean Force

The mean force can also be obtained by differentiation of the pmf according to Eq. 8. The pmf acting between the two macroions can be directly obtained by sampling the frequency of the macroion separation r, $P(r)$, from MC simulations of the model system. For example, the cylinder model can again be employed, but now having the macroions constrained to only the z-axis and allowing their z-coordinates to vary. From $P(r)$ the pmf becomes

$$U^{pmf}(r)/kT = -\ln\left[\frac{P(r)}{P(r \to \infty)}\right] . \tag{17}$$

In practice, $P(L_{cyl}/2)$ is used to normalize $U^{pmf}(r)$, i.e., $U^{pmf}(L_{cyl}/2) \equiv 0$ is assigned. In regimes where high electrostatic coupling appears, cluster trial moves (Sect. 6.3.3) are strongly advocated to enhance the simulation efficiency. When the pmf displays large variation ($> 5kT$), weighting functions can be introduced to obtain a more uniform sampling and hence an improved description of $P(r)$.

The accuracy of the pmf can be predicted according to

$$\sigma\left(U^{pmf}(r)\right) \approx \frac{1}{\sqrt{\langle N_{hit,i} \rangle}} , \tag{18}$$

where $\langle N_{hit,i} \rangle$ is the average number of hits in histogram i representing the distance r. With a uniform sampling, we have $\langle N_{hit,i} \rangle = N_{aqu}/N_{hist}$ with N_{aqu} being the number of data acquisitions and N_{hist} the number of histograms representing $P(r)$.

5.4 Comparison of Different Approaches

Which of the approaches should be selected to evaluate the mean force and/or the pmf between two macroions mediated by the small ions? In what follows, some characteristics of the three approaches will be compared and the complexity of extending them to other systems will be discussed.

Mean forces and their uncertainties have been calculated using the two mean-field approaches for Systems I–IV. The optimal Δr and Δh for the shorter separation given in Table 3 was used. The choice of Δr and Δh is important, since the uncertainty of F is dominated by the $F_{\rm hs}$ and $F_{\rm ideal}$ terms, respectively. Data for the two approaches were obtained from the same simulations.

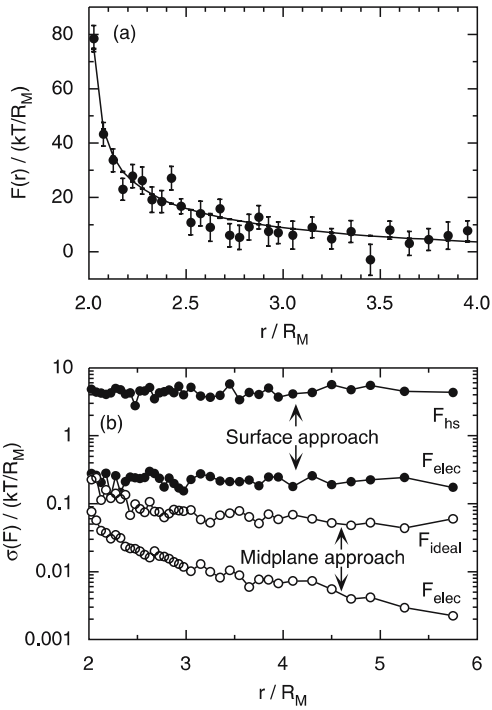

Fig. 8 **a** Mean force as a function of the macroion separation for System II evaluated using the surface approach (*filled symbols with error bars*) and the midplane approach (*solid curve with nearly invisible error bars*) and **b** uncertainty of the hard-sphere and electrostatic contributions of the surface approach (*filled circles*) and of the ideal and electrostatic contributions of the midplane approach (*open circles*). $R_{\rm cyl} = 4R_{\rm M}$, $L_{\rm cyl} = 12R_{\rm M}$, and $N_{\rm pass} = 10^6$. The uncertainty estimates are based on a division of the simulations into 20 blocks

Figure 8 shows the mean force and the uncertainties of the force components for System II as a function of the macroion separation. First, we conclude that the two approaches predict the same mean force within the statistical uncertainties. Second, the mean force calculated using the surface approach (Fig. 8a, filled circles) displays considerable scattering at all separations, whereas the corresponding data using the midplane approach (Fig. 8a, solid curve) are smooth (uncertainty bars are smaller than the thickness of the line). Regarding the surface approach, the uncertainty in F arises entirely from the uncertainty in the determination of F_{hs} (cf. curves with filled circles in Fig. 8b). In the midplane approach, the uncertainty in the determination of F_{ideal} dominates the uncertainty in F (cf. curves with open circles in Fig. 8b). A similar difference in the accuracy of the two approaches for highly coupled systems has been reported [28].

Representative uncertainties of the force components for Systems I–IV using the two approaches are summarized in Table 4. In all cases, the uncertainty in F is dominated by the uncertainty in F_{hs} (surface approach) and in F_{ideal} (midplane approach). The uncertainties are smallest for System I, and the performance of the two approaches is similar for this system. Also for System IV, the performance of the two approaches is comparable. For Systems II and III, however, the surface approach provides F with ≈ 50 times higher precision, corresponding to 2500 shorter simulation for similar accuracy. A comparison of the results for Systems II and IV shows a 25-fold difference in the uncertainty in F_{hs} of the surface approach (Table 4), despite the fact that the mean forces are not too dissimilar (Fig. 5). This large difference in the uncertainties is attributed to the larger inhomogeneity of the distribution of the small ions near the macroions occurring at the larger electrostatic coupling in System II (Fig. 4).

The accuracy of simulated pmf's for Systems I and III has also been examined. Table 5 shows that Eq. 18 provides a reasonable estimation of the

Table 4 Mean-force calculations: uncertainties in F_{elec} and F_{hs} for the surface approach and in F_{elec} and F_{ideal} for the midplane approach in kT/R_{M} units[a]

System	Surface approach[b]		Midplane approach[c]	
	$\sigma(F_{\mathrm{elec}})$	$\sigma(F_{\mathrm{hs}})$	$\sigma(F_{\mathrm{elec}})$	$\sigma(F_{\mathrm{ideal}})$
I	0.02	0.05	0.01	0.03
II	0.2	5	0.02	0.1
III	0.2	2	0.02	0.03
IV	0.06	0.2	0.03	0.1

[a] $R_{\mathrm{cyl}} = 4R_{\mathrm{M}}$, $L_{\mathrm{cyl}} = 12R_{\mathrm{M}}$, and $N_{\mathrm{pass}} = 10^6$
[b] optimal Δr according to Table 3
[c] optimal Δh according to Table 3

Table 5 Potential of mean force calculations: uncertainty in U^{pmf} in units of kT from simulations in which macroion separations are sampled[a]

	Eq. 18 [b]	System I	System III
$\sigma(U^{\mathrm{pmf}})$	0.005	0.005	0.015

[a] $R_{\mathrm{cyl}} = 4R_{\mathrm{M}}$, $L_{\mathrm{cyl}} = 12R_{\mathrm{M}}$, cluster trial displacements with $\Delta_{\mathrm{cyl}} = R_{\mathrm{M}}$, and $N_{\mathrm{pass}} = 10^8$. A pmf simulation with $N_{\mathrm{pass}} = 10^8$ involves about the same CPU time as 80 mean-force simulations comprising $N_{\mathrm{pass}} = 10^6$ passes.

[b] Data acquisition was made every 10 pass using the interval $2R_{\mathrm{M}}$–$12R_{\mathrm{M}}$ with a $0.05R_{\mathrm{M}}$-spacing (as used in the mean force calculations at short macroion separation), wich gives $N_{\mathrm{aqu}} = 10^7$ and $N_{\mathrm{hist}} = 200$ histograms resulting in $\langle N_{\mathrm{hist},i} \rangle = 5 \times 10^4$

uncertainty. Furthermore, for comparable computer time, the pmf approach provides more accurate results for System I and results of similar accuracy for System III as compared to the midplane mean-force approach (cf. data in Tables 4 and 5). For simplicity, the uncertainties in F and U^{pmf} have been compared directly without invoking any numerical derivation or integration, a reasonable procedure as long as lengths are given in numbers of order of unity.

5.5 Practical Guidelines

The two approaches to simulate the mean force and the one to simulate the pmf possess different characteristics in terms of (A) efficiency, (B) ease of use, and (C) extension to systems with more complex particles.

(A) Ample numerical evidences with theoretical understanding show that the mean-force determination using the midplane approach is more efficient than the surface approach as the accumulation of the small ions near the macroions becomes large. A direct determination of the pmf by a macroion-separation sampling provides results with similar or slightly higher efficiency than the midplane mean-force approach.

(B) The two mean-force approaches require optimization of parameters controlling the determination the hard-sphere and the ideal contributions, respectively. In the macroion-separation sampling, no such optimization is present. However, for electrostatically strongly coupled systems, cluster trial displacements are required and a weighting function could be needed to obtain an approximately uniform sampling.

(C) If the small ions contain a hard-sphere radius and/or if chain molecules are present, Eqs. 9–11 and 17 remain unchanged, whereas Eq. 13 should be extended with terms describing the transfer of linear moment across the midplane of these ions and/or bonding force across the midplane. Finally, if the macroion possesses a nonspherical symmetry, orientational av-

eraging could be applied in all three approaches to get an orientationally averaged mean force and pmf.

Hence, all three approaches have their own merits. If (i) the mean force is the primary target, (ii) the interaction is needed for only a limited interval, or (iii) a separation into different force components is desired, a mean-force sampling using the midplane approach is the approach of choice. However, if the electrostatic coupling is weak and, in particular, for systems with more complex composition, the surface approach could be an alternative. The approach involving macroion-separation sampling has its strength when full interaction curves are needed, in particular if the macroion displacements can be made efficiently.

6
Periodic System

As mentioned in Sect. 2.2, for bulk simulations, systems are normally made periodic to avoid surface effects. Nevertheless, large systems may still be needed to cope with long spatial correlation lengths and/or long-range interactions, and a number of CPU time-saving features are often applied to make such simulations feasible.

6.1
Potential Truncation

Typically, 90% or more of the computational time is devoted to calculations of the interparticle distances and potentials/forces. Substantial savings would be made if the interactions between a particle and only its neighboring particles needed to be considered. For systems with short-range interactions only, such a truncation of the interactions is straightforward. However, simulation of ionic systems with long-range Coulomb interactions requires special considerations. The consequences of truncating the Coulomb interaction in simple electrolyte and molten salt systems with periodic systems was addressed in the 1970s [118, 119].

6.1.1
General

The simplest truncation of the interparticle interaction is to make a spherical cutoff (SC) at the particle separation R_{cut}. For a Coulomb system, this approach suffers from the observation that the net charge of the central particle and the other particles inside the cutoff region is not necessarily zero, which creates undesired artifacts (Sect. 6.1.2).

A slightly more sophisticated method is the so-called "minimum-image" (MI) convention, in which all interactions between the central particle and its nearest image of all other particles are included. This is a significant improvement over a SC, since the central particle and its interacting particles constitute an electroneutral entity. The MI convention has frequently been applied to simulate simple electrolytes and, to some extent, weakly asymmetric electrolytes. However, it suffers from the same principal deficiencies as a SC, and the MI convention becomes insufficient for highly asymmetric electrolytes.

The Ewald summation is a well-suited method to handle highly asymmetric electrolytes. Due to the availability of efficient numerical implementations of the Ewald summation, it is also frequently used to simulate systems with a weak Coulomb coupling where the MI would work. Apart from the Ewald summation, multipole expansion techniques have also been employed to simulate large Coulomb systems.

6.1.2 Comparison of Different Methods

The SC, the MI convention, and the Ewald summation have been applied to Systems I and II. Figure 9 displays macroion–macroion RDFs for System I employing the three different boundary conditions.

With the SC, there is an extremely large distortion of $g_{MM}(r)$ at $r = R_{cut}$ (Fig. 9, solid curve; the arrow denotes the location of R_{cut}). Right below R_{cut} is

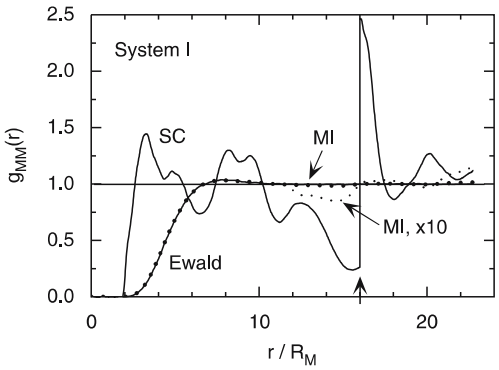

Fig. 9 Macroion–macroion radial distribution functions for System I using a spherical cutoff at $R_{cut} = 16R_M$ (SC, *full curve*), the minimum image convention (MI, *dotted curve*), and a magnification according $10(g_{MM}(r) - 1) + 1$ for $r \geq 12R_M$ (MI, $\times 10$, *dotted curve*), and the Ewald summation (Ewald, *full curve*). The RDFs have been analyzed beyond $r = L/2$ (*arrow*) by employing the appropriate volume elements in the normalization. The *thin solid line* corresponds to a uniform distribution. $N_M = 80$, $L = 32.224R_M$, and $N_{pass} = 5 \times 10^4$

a deficiency, and right beyond R_{cut} is an excess of other macroions. As previously analyzed [118], the cutoff leads to a too attractive potential beyond R_{cut}, giving rise to an enhanced local density by primarily transferring macroions from the region $r < R_{cut}$ to $r > R_{cut}$ with a concomitant reduction of the potential energy of the system. This distortion propagates from R_{cut} to both shorter and longer distances, making the SC obviously completely useless. In the MI convention, a given particle interacts with the nearest image of all other particles, making the distortion less drastic (Fig. 9, dotted curve labeled "MI"). Nevertheless, a tenfold magnification of the deviation from $g_{MM}(r) = 1$ around $r = L/2$ displays a reduced density at $r < L/2$ and an enhanced one in the diagonal directions at $r \geq 2^{1/2}(L/2)$. Obviously, macroions are depleted from the box sides and accumulated far away in the diagonal directions, also leading to a reduction of the potential energy of the system. Finally, with the Ewald summation no such irregularities are found, and this appraoch will be scrutinized in more detail below.

For the similar investigation of System II, which possesses much stronger direct macroion–macroion interaction, the artifacts of the SC or the MI convention become even more severe. Moreover, the fraction of accepted trial moves becomes very small, making the simulations practically nonergodic.

Hence, these observations agree with previous findings that (i) simulation of simple electrolytes can be performed with the MI convention but hardly with a SC and (ii) the MI convention can typically be employed for simulations of, at most, slightly asymmetric electrolytes [103, 104].

6.2 Ewald Summation

Since the Ewald summation is a well-suited and frequently used method to treat the interactions in strongly coupled Coulomb systems, a brief background will be given and its system size dependence and truncation error will be examined. Finally, some practical guidelines will be provided.

6.2.1 General

The Ewald summation is a method to evaluate the potential energy of a periodic system of infinite extension, originally designed to obtain the potential energies of ionic crystals [59]. When applied to simulations of liquids, the system is made infinite by replicating the primary simulation box in (here) three dimensions (Fig. 10). Hence, the system is disordered on some length scale smaller than L but periodic for longer distances, which introduces artifacts. For such a truly periodic system containing charges, the expression of

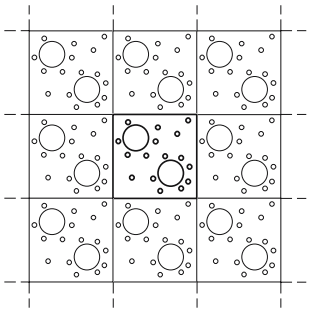

Fig. 10 Schematic illustration showing the primary (*central*) simulation box containing charged macroions and charged counterions as well as an infinite replication of the primary box

the total Coulomb energy becomes

$$U = \frac{1}{4\pi\varepsilon_0\varepsilon_r} \sum_{i<j} {\sum_{\boldsymbol{m}\in Z^3}}' \frac{q_i q_j}{|\boldsymbol{r}_{ij} + \boldsymbol{m}L|}, \tag{19}$$

where $q_i = Z_i e$ and $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$, with $\boldsymbol{r}_i$ denoting the location of particle i, the sum over $\boldsymbol{m}$ reflects the different periodic images, and the prime symbolizes that the term $\boldsymbol{m} = 0$ should be omitted for $i = j$.

By superimposing each charged particle with a Gaussian charge distribution of width $\alpha^{-1/2}$ and of opposite sign, the total Coulomb interaction can be reformulated according to [59, 64]

$$\begin{aligned} {\sum_{i<j}}' {\sum_{\boldsymbol{m}\in Z^3}}' \frac{q_i q_j}{|\boldsymbol{r}_{ij} + \boldsymbol{m}L|} = {\sum_{i<j}}' {\sum_{\boldsymbol{m}\in Z^3}}' \frac{q_i q_j}{|\boldsymbol{r}_{ij} + \boldsymbol{m}L|} \operatorname{erfc}(\alpha|\boldsymbol{r}_{ij} + \boldsymbol{m}L|) \\ + \frac{2\pi}{L^3} \sum_{\boldsymbol{k}\neq 0} k^{-2} \exp(-k^2/4\alpha^2) \left| \sum_j q_j \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j) \right|^2 \\ - \frac{\alpha}{\sqrt{\pi}} \sum_j q_j^2 \\ + \frac{2\pi}{(1+2\varepsilon')L^3} \left| \sum_j q_j \boldsymbol{r}_j \right|^2, \end{aligned} \tag{20}$$

where erfc denotes the complementary error function, $\boldsymbol{k} \equiv (2\pi/L)\boldsymbol{n}$ with $\boldsymbol{n} \equiv (n_x, n_y, n_z)$ being a box translation vector and n_x, n_y, and n_z integer numbers. The first term on the right-hand side of Eq. 20 denotes the interaction among the charged particles (now screened) in the real space, the second term the interaction among the compensating Gaussian charge distributions expressed in the reciprocal space, and the third term a correction for the

self-energy of the Gaussian charge distributions. In the last term of Eq. 20, ε' denotes the permittivity of the surrounding medium, and the expression of this term depends on the summation order (here spherical) over the replicas in the reciprocal space [120]. If the surrounding medium is conducting (ε' being infinite), this term vanishes. There are indications that the conducting condition is the preferred one for Coulomb [108, 121] and dipolar [122] systems.

Several algorithms for implementing the Ewald summation are available. The straightforward approach of using a double loop for finding pairs of interacting particles in the real space and a single particle loop in the reciprocal space has a performance of $O(N^{3/2})$ [123], whereas more complex particle-mesh Ewald implementations [124, 125] display the weaker $O(N \ln N)$ dependence, where N is the number of particles. The crossover of these two approaches depends on the particular system, the degree of optimization of the code, and the hardware, but appears about $O(10^4)$ for Coulomb systems [126, 127]. More details, including other mesh methods and multipole expansions, are found in the chapters by C. Holm and by J.J. Weis and D. Levesque.

In the implementation of the Ewald summation according to Eq. 20, the value of the potential energy is controlled by three parameters: α, the upper limit of $\boldsymbol{m}$ ($\boldsymbol{m}_{\mathrm{cut}}$), and the upper limit of $\boldsymbol{n}$ ($\boldsymbol{n}_{\mathrm{cut}}$). At equal truncation error in the two spaces, the summation in the real space is often limited to interactions involving only the nearest image ($\boldsymbol{m} = 0$), and consequently a spherical cutoff distance $R_{\mathrm{cut}} \leq L/2$ in the real space can be applied. Moreover, the number of replicas in reciprocal space can be reduced by applying a spherical cutoff of $\boldsymbol{n}$ according to $|\boldsymbol{n}| \leq n_{\mathrm{cut}}$.

Kolafa and Perram [128] have analyzed the truncation error of the Ewald summation. Assuming uncorrelated positions beyond R_{cut}, the statistical error of the real-space term of Eq. 20 becomes, in our notation,

$$\frac{\delta_{\mathrm{Ureal}}}{NkT} \approx R_{\mathrm{M}} \Gamma_{\mathrm{II}} Z_{\mathrm{r}} \left(\frac{R_{\mathrm{cut}}}{2L^3}\right)^{1/2} \left(\frac{1}{\alpha R_{\mathrm{cut}}}\right)^2 \exp\left[-(\alpha R_{\mathrm{cut}})^2\right], \tag{21}$$

whereas the error in the reciprocal space becomes

$$\frac{\delta_{\mathrm{Urec}}}{NkT} \approx R_{\mathrm{M}} \Gamma_{\mathrm{II}} Z_{\mathrm{r}} \frac{n_{\mathrm{cut}}}{L} \left(\frac{\alpha L}{\pi n_{\mathrm{cut}}}\right)^2 \exp\left[-\left(\frac{\pi n_{\mathrm{cut}}}{\alpha L}\right)^2\right] \left\{1 + \frac{1}{\alpha L n_{\mathrm{cut}}^{1/2}}\right\}, \tag{22}$$

where the first term after expanding the curly brackets denotes a systematic and configurationally independent error and the second term a statistical error where uncorrelated positions beyond n_{cut} again have been assumed. By assigning a largest tolerated error δ_{tol} of the real-space and reciprocal-space terms, R_{cut} and n_{cut} become an implicit function of δ_{tol} and α through Eqs. 21 and 22. As we will shortly see, $\delta_{\mathrm{Urec}} \leq \delta_{\mathrm{Ureal}} = \delta_{\mathrm{tol}}$ is obtained, since R_{cut} is as a continuous and n_{cut} a discontinuous variable. Finally, α is selected

to minimize the overall computational load, and at this stage the condition $R_{\mathrm{cut}} \leq L/2$ has to be checked.

This analysis can be simplified by noticing that

$$\frac{\delta_{\mathrm{Ureal}}}{NkT} \propto \left(\frac{1}{\alpha R_{\mathrm{cut}}}\right)^2 \exp\left[-(\alpha R_{\mathrm{cut}})^2\right], \tag{23}$$

$$\frac{\delta_{\mathrm{Urec}}}{NkT} \propto \left(\frac{\alpha L}{\pi n_{\mathrm{cut}}}\right)^2 \exp\left[-\left(\frac{\pi n_{\mathrm{cut}}}{\alpha L}\right)^2\right], \tag{24}$$

where (i) the two omitted prefactors are independent of α and only weakly dependent on R_{cut} and n_{cut}, respectively, and (ii) the statistical error in Eq. 22 is neglected. For realistic values of α and n_{cut}, we have $1/(\alpha L n_{\mathrm{cut}}^{1/2}) < 0.1$. For example, $\alpha = 6/L$ and $n_{\mathrm{cut}} = 6$ gives $1/(\alpha L n_{\mathrm{cut}}^{1/2}) \approx 0.07$. The assignment of equal truncation errors in the two spaces and the definition $s \equiv \alpha R_{\mathrm{cut}} = \pi n_{\mathrm{cut}}/\alpha L$ now lead to

$$\frac{\delta_{\mathrm{Ureal}}}{NkT} \approx \frac{\delta_{\mathrm{Urec}}}{NkT} \propto f(s) \equiv \frac{1}{s^2} \exp\left(-s^2\right). \tag{25}$$

Hence, the cutoff distance of the pair interactions in real space R_{cut} and the cutoff in reciprocal space n_{cut} become explicit functions of s (the variable now controlling the accuracy) and α according to [65, 108]

$$R_{\mathrm{cut}} = \frac{s}{\alpha}, \tag{26}$$

$$n_{\mathrm{cut}} = 1 + \left[\frac{s\alpha L}{\pi}\right], \tag{27}$$

where in Eq. 27 [...] denotes a truncation to an integer number.

6.2.2 System Size Convergence

As was mentioned, the use of the Ewald summation implies an imposed periodic structure on the fluid under investigation. The effect of this imposed periodicity is conventionally examined by performing simulations using different system sizes. It is conjectured that the effect of the periodicity on the fluid properties should asymptotically vanish in the thermodynamic limit.

Such system size dependences have been investigated for Systems I–IV employing systems containing $N_{\mathrm{M}} = 20$, 40, and 80 macroions in the primary box using the Ewald summation with conducting boundary conditions. Table 4 provides the potential energies and pressures, whereas Fig. 11 shows the macroion–macroion RDF for System IV, the system displaying the largest macroion structuring (Fig. 6a).

Despite extended simulations, the thermodynamic data in Table 6 show hardly any significant system size dependence of the potential energy or the pressure. Similarly, only very small effects on the macroion–macroion RDF

Table 6 Reduced energy and reduced pressure at different numbers of macroions N_M using the Ewald summation with conducting boundary conditions[a]

System	N_M	$U/N_M kT$	pV/NkT
I	20	–0.7729(4)	0.820(1)
	40	–0.7719(4)	0.821(1)
	80	–0.7721(6)	0.821(2)
II	20	–13.2679(4)	0.234(4)
	40	–13.2668(4)	0.234(4)
	80	–13.2680(8)	0.219(7)
III	20	–30.293(2)	–
	40	–30.271(1)	–
	80	–30.274(2)	–
IV	20	–0.7994(2)	0.8115(5)
	40	0.7996(2)	0.8115(4)
	80	0.7996(3)	0.8113(6)

[a] $N_{pass} = 4 \times 10^5$ ($N_M = 20$), 2×10^5 ($N_M = 40$), and 5×10^4 ($N_M = 80$)

can be seen in Fig. 11. With $N_M = 20$, the magnitude of the first maximum is slightly exaggerated, whereas the RDF with $N_M = 40$ displays only minute differences as compared to that with $N_M = 80$ for the distances that are in common. Hence, even for systems containing macroions with 80 elementary charges and having an interacting energy $> 500kT$ with neighboring macroions positioned in the first coordination shell ($r \approx 8R_M$), the Ewald summation with as few as $N_M = 20$ macroions provides results very close to

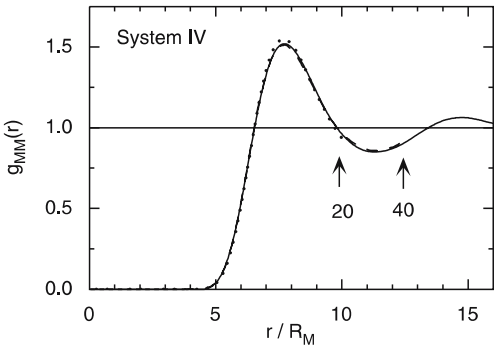

Fig. 11 Macroion–macroion radial distribution functions for System IV with $N_M = 20$ (*dotted curve*), $N_M = 40$ (*dashed curve*), and $N_M = 80$ (*solid curve*) macroions. The two *arrows* indicate the ends of the calculated distribution functions for $N_M = 20$ and $N_M = 40$, respectively. The *thin solid line* corresponds to a uniform distribution. $N_{pass} = 4 \times 10^5$ ($N_M = 20$), 2×10^5 ($N_M = 40$), and 5×10^4 ($N_M = 80$)

the thermodynamic limit. Noticeably, the predicted structure up to $L/2$ is essentially correct, although the spatial correlations have not yet decayed. The common recommendation to select a system size, which is able to accommodate the longest correlation length of the system, can obviously be relaxed to some extent.

In other words, the Ewald summation displays here a rapid system size convergence, making simulations with a fairly limited number of particles feasible. This is important for the examination of asymmetric electrolytes, where each macroion is accompanied by a large number of counterions and in particular when salt is added.

6.2.3 Summation Convergence

The role of the different parameters controlling the Ewald summation will now be considered in more detail. A single equilibrium configuration of System IV with $N_M = 80$ macroions will be used, but the outcome does not critically depend on the choice of system or configuration. Values of the potential energy for different truncations of the Ewald summation will be compared with essentially the exact value, the latter obtained from an Ewald summation with large values of R_{cut} and n_{cut}.

6.2.3.1 Variable α at fixed R_{cut} and n_{cut}

As mentioned in Sect. 6.2.1 (cf. Eq. 20), the accuracy of the contributions from the real and reciprocal spaces to the potential energy depends on α and R_{cut} and on α and n_{cut}, respectively. Figure 12 shows the absolute value of the difference between the reduced potential energy from the Ewald summation and the exact one as a function of α for *fixed* R_{cut} and n_{cut} using an equilibrium configuration (filled symbols) and a random one but avoiding hard-sphere overlap (open symbols). The error estimates given by Eqs. 21 and 22 are also shown (curves).

An increasing error in the potential energy appears as α is reduced at $\alpha < 0.3R_M^{-1}$ *and* as α is increased at $\alpha > 0.3R_M^{-1}$ (spheres). The first effect is a consequence of the $\mathrm{erfc}(\alpha r)/r$ function in the real-space term of Eq. 20, which becomes more long-range as α is reduced, whereas the truncation error at large α arises from the reciprocal-space term of Eq. 20 with the $\exp(-1/\alpha^2)$ dependence. At intermediate α, the truncation error displays a minimum due to an accurate summation of both terms, the minimum appearing at $\alpha \approx 0.3R_M^{-1}$ and possessing a negligible error $[O(10^{-10}NkT)]$. The size of the error at the minimum is controlled by R_{cut} and n_{cut} collectively.

The truncation errors for the equilibrium and random configurations are similar in the real space and virtually identical in the reciprocal space,

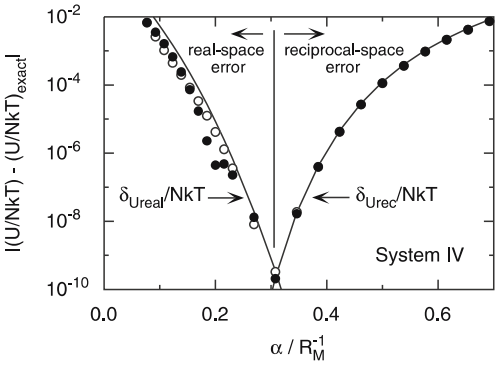

Fig. 12 Absolute truncation error of the reduced potential energy from the Ewald summation as a function of α at $R_{\text{cut}} = 13.0R_{\text{M}}$ and $n_{\text{cut}} = 14$ using an equilibrium (*filled spheres*) and random (*open spheres*) configuration of System IV. For $\alpha > 0.38R_{\text{M}}^{-1}$, the symbols for the equilibrium and random configurations are indistinguishable. $N_{\text{M}} = 80$ and $L = 32.224R_{\text{M}}$. The hard-sphere radii were retained when generating the random configuration. The corresponding estimated truncation errors of the real-space and reciprocal-space terms of the reduced potential energy according to Eqs. 21 and 22, respectively, are also shown (*solid curves*)

demonstrating that the numerical errors are only weakly dependent on the configuration. A slight irregularity of the real-space error for the equilibrium configuration (filled circles) appears at $\alpha \approx 0.2R_{\text{M}}^{-1}$, and at this value of α the real-space error changes sign. This feature is attributed to the spatial correlations among the particles, since the corresponding result for the random configuration (open circles) is smooth and the real-space error does not change sign. The almost perfect agreement in the reciprocal space is rationalized by the fact that the systematic and configurationally independent term dominates over the configurationally dependent one.

Figure 12 also demonstrates a very good agreement between the analytic error estimates (Eqs. 21 and 22) and the numerical data over eight decades. In the real space, the estimated error is overestimated up to a factor of 2, whereas the estimated error in the reciprocal space agrees excellently with the numerical data. Hence, these estimates accurately describe the actual truncation errors, and the practical use of such estimates will now be discussed.

6.2.3.2 R_{cut} and n_{cut} as Dependent Variables

The perspective on the variables controlling the Ewald summation will now be changed. The variables δ_{tol} and α will now be considered as the independent ones with R_{cut} and n_{cut} being the dependent ones through Eqs. 21 and 22, *or*, in the simplified analysis, s and α are the independent variables with R_{cut} and n_{cut} determined through Eqs. 26 and 27.

Figure 13 displays absolute values of the truncation error of the reduced potential energy as a function of $\delta_{\mathrm{tol}}/NkT$ at a fixed α and as a function of α at a fixed δ_{tol}, in both cases with R_{cut} and n_{cut} determined through Eqs. 21 and 22 using an equilibrium configuration of System IV (spheres). The real-space $\delta_{\mathrm{Ureal}}/NkT$ and reciprocal-space $\delta_{\mathrm{Urec}}/NkT$ error estimates of the potential energy given by Eqs. 21 and 22, respectively, are also shown in Fig. 13 (solid curves). First, as has already been alluded to, $\delta_{\mathrm{Urec}} \leq \delta_{\mathrm{Ureal}} = \delta_{\mathrm{tol}}$, and hence the truncation error is larger in the real space than in the reciprocal one. Second, the observed truncation errors deviate normally at most by a factor of 3 from δ_{tol}. Occasionally, the observed error becomes smaller, which appears when it changes sign, e.g., at $\delta_{\mathrm{tol}}/NkT = 5 \times 10^{-4}$ (Fig. 13a) and $\alpha/R_{\mathrm{M}}^{-1} = 0.28$ (Fig. 13b). Third, the observed errors do not display any systematic drift as α is changed (Fig. 13b).

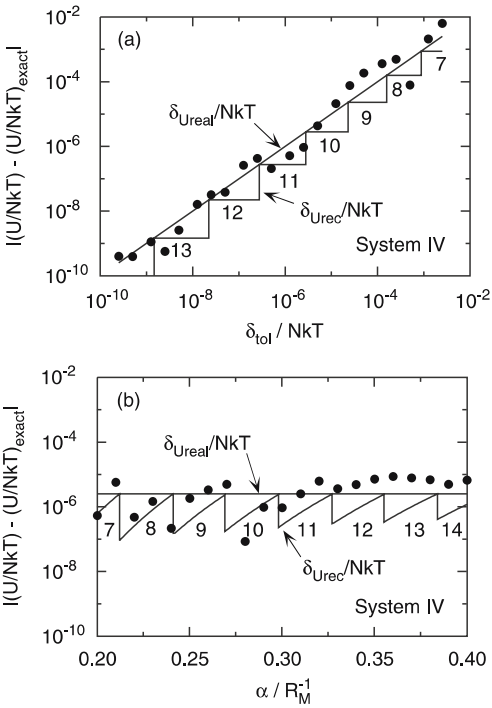

Fig. 13 Absolute truncation error of the reduced potential energy of the Ewald summation as a function of **a** the potential energy tolerance $\delta_{\mathrm{tol}}/NkT$ at $\alpha = 0.3R_{\mathrm{M}}^{-1}$ and **b** α at $\delta_{\mathrm{tol}}/NkT = 2.5 \times 10^{-6}$ with R_{cut} and n_{cut} evaluated using Eqs. 21 and 21 and the same equilibrium configuration of System IV as in Fig. 12 (*symbols*). The corresponding estimated truncation errors of the real-space and reciprocal-space terms of the reduced potential energy according to Eqs. 21 and 22, respectively (*solid curves*), and the values of n_{cut} are also shown

The corresponding data for the simplified error analysis are given in Fig. 14, where $\delta_{\mathrm{tol}}/NkT$ is replaced by $f(s) \equiv \exp(s^2)/s^2$. First, we notice that $\delta_{\mathrm{Ureal}}/NkT$ and $\delta_{\mathrm{Urec}}/NkT$ are essentially proportional to $f(s)$ (nearly unit and zero slope in Fig. 14a,b, respectively), confirming that the prefactors omitted in Eqs. 23 and 24 are only weakly dependent on R_{cut} and n_{cut}. Second, $\delta_{\mathrm{Urec}} \approx 10\delta_{\mathrm{Ureal}}$, consistent with the numerical values of the prefactors. Third, Fig. 14a shows that observed truncation errors (i) display a stepwise behavior and follow closely δ_{Urec}, consistent with $\delta_{\mathrm{Ureal}} < \delta_{\mathrm{Urec}}$, and (ii) on a coarser scale are virtually proportional to $f(s)$ with a prefactor of 0.4. Fourth, Fig. 14b shows that the observed truncation error remains essentially constant in the investigated α interval. The limit $\alpha = 0.2R_{\mathrm{M}}^{-1}$ corresponds to $(R_{\mathrm{cut}},\ n_{\mathrm{cut}}) = (15R_{\mathrm{M}}, 7)$, whereas $\alpha = 0.4R_{\mathrm{M}}^{-1}$ to $(R_{\mathrm{cut}}, n_{\mathrm{cut}}) = (6.67R_{\mathrm{M}}, 13)$. The serrated pattern appearing is also a consequence of the discrete changes of

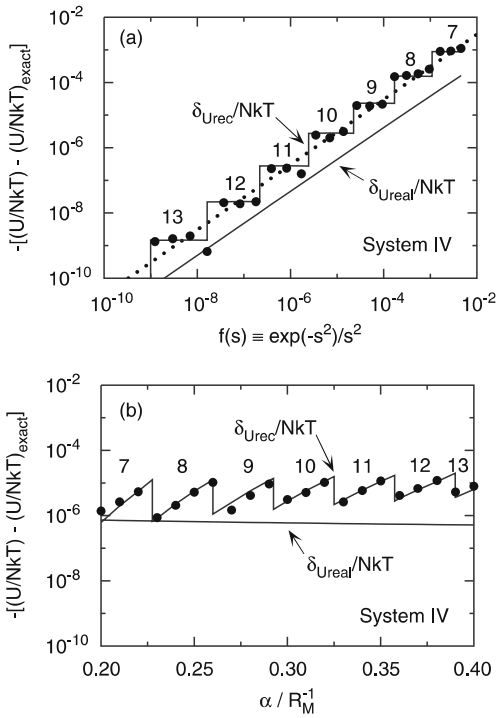

Fig. 14 Absolute truncation error of the reduced potential energy of the Ewald summation as a function of **a** $f(s) \equiv \exp(s^{-2})/s^{-2}$ at $\alpha = 0.3R_{\mathrm{M}}^{-1}$ and **b** α at $s = 3.0$ from simulation with R_{cut} and n_{cut} evaluated according to Eqs. 26 and 27, respectively, and the same equilibrium configuration of System IV as in Fig. 12 (*symbols*). The corresponding estimated truncation errors of the real-space and reciprocal-space terms of the reduced potential energy according to Eqs. 21 and 22, respectively (*solid curves*), and the values of n_{cut} are also shown. In **a**, the function $g(s) = 0.4f(s)$ is given (*dotted line*)

n_{cut}. Nevertheless, the observed truncation error displays only a very weak systematic increase over the α-interval, the magnitude of the increase being less than the irregularity originating from the discrete changes of n_{cut}.

To summarize: (A) With the full error analysis $\delta_{Urec} \leq \delta_{Ureal}$ due to the discrete variation of n_{cut}, whereas with the simplified analysis $\delta_{Ureal} \leq \delta_{Urec}$ due to different values of omitted prefactors. (B) Typical ratios of the real-space and reciprocal-space truncation errors are a factor of 5 for the full and 10 for the simplified analysis. The ratios are comparable to (i) the relative discontinuous change in δ_{Urec} (a factor of 10) as $n_{cut} \rightarrow n_{cut} \pm 1$ and (ii) the difference between the error estimate and the actual truncation error (a factor of 5). (C) The simplified error analysis (i) accurately describes the functional form of the truncation error of the potential energy and (ii) predicts R_{cut} and n_{cut}, which provides an essentially constant truncation error as α is varied. From (B) and (C) we conclude that the simplified error analysis provides the functional form of the truncation error and R_{cut} and n_{cut} with constant (but yet unknown) truncation error.

6.2.4 Optimization of Execution Time

It was shown in Figs. 13b and 14b that the accuracy of the Ewald summation is essentially independent of α provided R_{cut} and n_{cut} are selected according to Eqs. 21 and 22 or Eqs. 26 and 27, respectively. However, the execution time of the energy evaluations depends on α, and hence α could be selected to minimize the execution time [129].

Figure 15 illustrates how the execution time of the energy evaluation t_{tot} and its decomposition in the time for the evaluations in the real t_{real} and re-

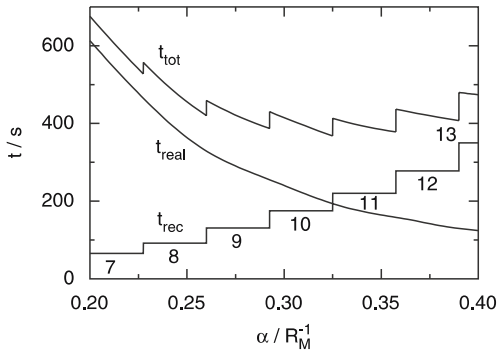

Fig. 15 Execution time of the potential energy in the real space t_{real}, in the reciprocal space t_{rec}, and their sum t_{tot} as a function of α at $s = 3.0$ with R_{cut} and n_{cut} evaluated according to Eqs. 26 and 27 from short MC simulations of System IV. The values of n_{cut} are also shown

ciprocal t_{rec} spaces can depend on α at fixed s. A small α leads to a large R_{cut} to maintain the required accuracy in the real-space term and t_{real} becomes large due to a large number of particle pairs to be evaluated, whereas a large α implies a large n_{cut} and t_{rec} becomes large due to the large number of $\boldsymbol{n}$-vectors. Moreover, the computational load in the reciprocal space varies in a stepwise pattern as proposed by Eq. 27, and the relevant values of n_{cut} are displayed in Fig. 15. The total time t_{tot} displays a weak minimum at $\alpha \approx 0.30 - 0.35 R_M^{-1}$, although a marked serrated pattern obtained from the stepwise behavior of t_{rec} is visible.

The shallowness and serrated behavior of t_{tot} make extensive optimization unnecessary. The precise location of this minimum depends on the relative computation times for the potential energy evaluations in the real and reciprocal spaces. These computation times can depend substantially on (i) the algorithms employed (type of neighbor list, direct or tabulated energy evaluation in the real space, implementation of the evaluation of the reciprocal term, single or parallel computation, etc.) and (ii) the architecture of the computer used (CPU computation speed, cache sizes, memory bandwidth, etc.), making such an optimization useful when changing code, compiler, or computer hardware.

6.2.5 Practical Guidelines

It has been shown that the Ewald summation is an excellent tool for handling Coulomb interactions in highly asymmetric electrolytes. In particular, (i) the system size convergence, (ii) the truncation error, and (iii) the selection of α for optimization of the CPU time have been discussed. In practice, issues (i) and (ii) are interrelated; values of R_{cut} and n_{cut} for a given truncation error depend on system size and simulation results with controlled uncertainties are required to assess the system size dependence.

In practice, when considering a new system the following protocol is recommended:

(A) Select an acceptable truncation error δ_{tol} (or s) of the Ewald summation. Typically, a truncation error of 10^{-3}–10^{-2} times the expected statistical precision of the potential energy is satisfactory.

(B) Select an initial system size, which at this stage could be fairly small. An assessment that the system size is sufficiently large will be made later.

(C) Determine R_{cut} and n_{cut} using the selected δ_{tol} (or s) and an initial value of α. The determination of the optimal value of α will be made later. Equilibrate and perform a production simulation.

(D) Increase the system size, determine new R_{cut} and n_{cut} using the selected δ_{tol} (or s) and a reasonable value of α, and equilibrate and perform a production simulation.

(E) Compare results for the different system sizes. If the effects of the finite system size are not acceptable, return to (D).

(F) Determine the optimal α by performing several short simulations starting from an equilibrated configuration and using different values α.

(G) Check the truncation error for the selected α, R_{cut}, and n_{cut} against a well-converged Ewald summation for one or some equilibrated configurations. If the actual truncation error is not acceptable, decrease δ_{tol} (or increase s) and return to (B).

Finally, it has been stressed that α should be considered as an independent variable with R_{cut} and n_{cut} as dependent ones. However, sometimes it is more convenient to consider R_{cut} as an independent variable with α and n_{cut} as the two dependent ones. That might physically be more intuitive and makes it easy to control that R_{cut} does not exceed L/2. Special considerations should, however, be made if R_{cut} also controls the cutoff of other non-Coulomb interaction terms.

6.3 Trial Displacements

Numerical averages calculated from simulations of finite lengths are subjected to statistical uncertainties. In this section, the focus will be on how the selection of trial displacements affects the *precision* (accuracy) of the mean values obtained from MC simulations (of finite lengths). After a general introduction, the single-particle displacement protocol will be examined, and thereafter two enhanced displacement procedures enabling simulation of highly coupled asymmetric electrolytes will be reviewed.

6.3.1 General

In practical use, the Metropolis MC method [64, 66, 130] provides successive configurations that are correlated. The statistical uncertainty of a property determined from Metropolis MC simulations depends both on (i) the length of the simulation, i.e., the number of configurations, and (ii) the correlation among the configurations. Consequently, the precision of a property can be improved by making the simulation longer and/or reducing the correlation between successive configurations. Appropriate selection of type and size of trial displacement can drastically reduce the correlation. In the simplest type of trial displacement, a single particle is moved at a time, implying that two successive configurations differ only by the position of one particle.

The effect of the correlation between successive configurations on a given property is conventionally expressed as a statistical efficiency [64] or by a closely related autocorrelation "time" of that property [64, 66]. The precision of different properties is affected differently by the correlation and

hence has different statistical efficiency and autocorrelation time. Generally, the potential energy converges relatively fast, whereas properties depending on collective structural rearrangements converge slowly.

NB: If the relevant part of the configurational space is not fully spanned, a statistical analysis may lead to a seemingly accurate determination of the mean value of a property. Nevertheless, the mean value could be erroneous. In those cases, the bottleneck of the mobility in the configurational space needs to be analyzed, and the transitions through this bottleneck have to be improved.

Simulations of solutions containing particles with strong electrostatic coupling, like highly charged colloids with counterions or solutions of macromolecules with opposite charges, are particularly challenging. In such systems, particles of opposite charge are spatially close to each other, which reduces the ability to create a new configuration that substantially differs from the previous one. In particular, with single-particle trial displacements, successive configurations become highly correlated.

In simulations of charged colloids using the primitive model of electrolytes, the small magnitude of the accumulated (total) macroion displacement constitutes a problem [103]. Therefore, an examination of the macroion root-mean-square (rms) displacement per MC pass, $\langle \Delta r_M^2 \rangle^{1/2}$, is particularly useful to (i) determine the optimal values of the parameters controlling the trial moves and (ii) compare the performance of different trial displacement procedures. A practical condition for an acceptable simulation of fluids is that the total macroion displacement should, at least, amount to the macroion spatial correlation length. For not too small N_{pass}, $\langle \Delta r_M^2 \rangle^{1/2}$ grows as $N_{pass}^{1/2}$ [109].

The performance of the different displacement procedures will be illustrated by using Systems I–IV with a primary simulation box containing $N_M = 20$ macroions and simulations comprising $N_{pass} = 10^4$ passes. This system size is sufficient for the present purpose.

6.3.2 Single-particle displacement

The conventional Metropolis MC procedure involves attempts to select individual particles subjected to trial displacements in a random fashion. Typically, a trial move involves translational displacements in the x-, y-, and z-directions of one particle with displacement lengths uniformly sampled from $[-\Delta/2, \Delta/2]$. This makes the underlying matrix of selecting configurations symmetric (equal probabilities of selecting the forward and backward transitions) [64]. The trial move is accepted with the probability

$$P_{acc} = \min\{1, \exp[-(U_{new} - U_{old})/kT]\}, \tag{28}$$

where U_{new} and U_{old} are the potential energies of the new and old configurations, respectively. The choice of a symmetric underlying matrix and the

acceptance probability given by Eq. 28 makes the detailed-balance condition fulfilled.

As mentioned, the translation displacement parameter Δ affects the statistical efficiency of simulated properties, and it is normally selected differently for different types of particles. For our asymmetric electrolyte, denote the macroion and small-ion displacement parameters by Δ_M and Δ_I, respectively. Due to the extensive accumulation of counterions near macroions, the approach with single-particle trial moves leads to a very restricted total macroion displacement and hence to poor statistics [103]. When a hard-core overlap is encountered, i.e., $r < R_i + R_j$, U_{new} becomes formally infinite (Eq. 1) and P_{acc} becomes zero. Most often $\langle \Delta r_I^2 \rangle^{1/2} \gg \langle \Delta r_M^2 \rangle^{1/2}$ for a broad range of Δ_I, making it only necessary to examine how $\langle \Delta r_M^2 \rangle^{1/2}$ depends on Δ_M for optimizing the trial displacements.

Figure 16 displays the acceptance ratio (fraction of accepted macroion trial moves) and $\langle \Delta r_M^2 \rangle^{1/2}$ for macroion trial displacements with Δ_M ranging from $10^{-4} R_M$ to $10 R_M$. Generally, at small Δ_M the acceptance ratio tends to unity, whereas as Δ_M is increased the acceptance ratio is reduced (Fig. 16a). The conventional rule of thumb, albeit with limited theoretical support, is to select a displacement parameter giving an acceptance ratio of ≈ 0.5 (dotted line). Using this criterion, the optimal macroion displacement parameter becomes $\Delta_M \approx 2R_M$ (System I), $\Delta_M \approx 0.2R_M$ (System IV), and $\Delta_M \approx 0.005R_M$ (Systems II and III). The 400-fold variation of optimal values of Δ_M demonstrates that the selection of Δ_M could be crucial.

The corresponding macroion rms displacement depends linearly on Δ_M at small Δ_M (Fig. 16b). In the limit where all trial displacements are accepted, $\langle \Delta r_M^2 \rangle^{1/2} = \Delta_M/2$ (dotted line), and indeed this limit is reached at sufficiently small Δ_M. For sufficiently dense systems, we expect $\langle \Delta r_M^2 \rangle^{1/2}$ to decrease at large Δ_M due to frequent hard-sphere overlaps. For three of the four systems, such a maximum in $\langle \Delta r_M^2 \rangle^{1/2}$ is observed for the Δ_M interval investigated. The maximum in $\langle \Delta r_M^2 \rangle^{1/2}$ appears at $\Delta_M \approx R_M$ (System IV) and $\Delta_M \approx 0.02R_M$ (Systems II and III). For Systems II–IV, the optimal Δ_M value (i) gives an acceptance ratio of ≈ 0.1 and (ii) is fourfold larger than those giving a 0.5 acceptance ratio.

For Systems I and IV, the magnitude of the maximal $\langle \Delta r_M^2 \rangle^{1/2} \geq 0.1R_M$. For a simulation involving $N_{pass} = 5 \times 10^4$ passes (cf. Sect. 4.3), the total macroion rms displacement becomes $\geq 20R_M$, being much longer than the macroion correlation length (cf. Fig. 6a). However, for Systems II and III, the maximal $\langle \Delta r_M^2 \rangle^{1/2}$ is $\approx 0.001R_M$, which, with the same simulation length, gives a total macroion rms displacement of $0.2R_M$, obviously completely insufficient. To obtain a total macroion rms displacement of $10R_M$, 2500 times longer simulation would be required.

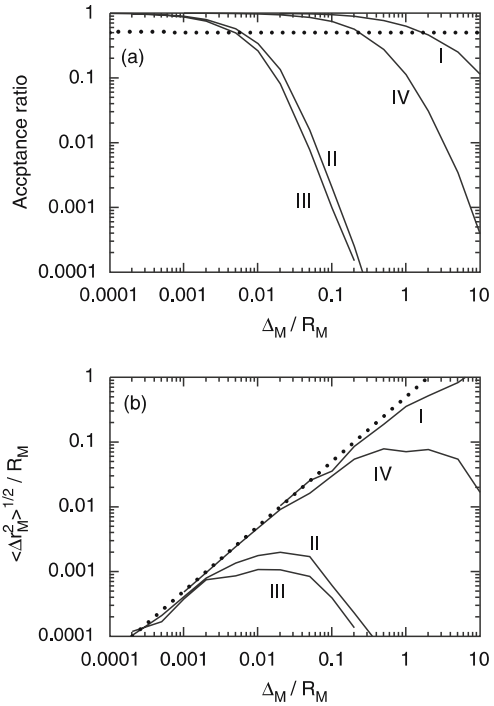

Fig. 16 **a** Macroion acceptance ratio and **b** macroion root-mean-square displacement per MC pass as a function of the macroion displacement parameter for indicated systems from simulations with single-particle trial moves. In **a**, the acceptance ratio equal to 0.5, and in **b**, the relation $\left\langle \Delta r_M^2 \right\rangle^{1/2} = \Delta_M/2$ are given (*dotted lines*). $N_M = 20$ and $N_{pass} = 10^4$

Figure 4 shows that the accumulation of counterions near macroions is much more prominent in Systems II and III as compared to the other two systems. The counterion densities at contact $\rho_I g_{MI}(R_M)$ are 0.021, 7.0, 9.2, and $0.15R_M^{-3}$ for Systems I–IV, respectively. Hence, the systems can be ordered with increasing density at contact according to I $\ll$ IV $\ll$ II < III, the sequence correlating excellently with the displacement analysis. Hence, this density analysis supports the view that the rejection of the macroion trial displacement originates from the large amount of nearby counterions.

6.3.3 Cluster Displacement

There is obviously a need to use more-advanced displacement procedures when a strong macroion–counterion coupling is present. The simulation efficiency can be improved dramatically by applying a displacement procedure where strongly coupled particles are displaced simultaneously, referred to here as a cluster move. Such ideas have been applied to MC simulations of

electrostatically coupled systems [60] and appeared elsewhere (see [65] for a review and further references).

In the present context, a cluster is composed of a central macroion and neighboring small ions. Small ions located within the distance R_{cl} from the macroion center will belong to the cluster with a probability $p_{cl} \leq 1$ [109]. The trial move of the whole cluster *as one entity* is accepted with the probability

$$P_{acc} = \min\{1, (1 - p_{cl})^{n_{new}-n_{old}} \exp[-(U_{new} - U_{old})/kT]\}, \quad (29)$$

where n_{new} and n_{old} are the numbers of particles inside the cluster volume after and before the trial move, respectively, and U_{new} and U_{old} have the same meaning as in Eq. 28. Again assuming a symmetric underlying transition matrix, the additional factor prior to the Boltzmann weight in the expression of the acceptance probability (cf. Eqs. 28 and 29) preserves the detailed-balance condition [60]. The cluster displacement procedure is further illustrated in Fig. 17. Typically, a cluster move is performed when a macroion is selected, whereas a conventional single-particle displacement is made when a small ion is chosen. The efficiency of the cluster displacement algorithm depends on (i) the small-ion inclusion probability p_{cl}, (ii) the cluster radius R_{cl}, and (iii) the cluster displacement parameter Δ_{cl}. The optimal value of p_{cl} for highly charged macroions was found to be very close to unity [109]. To simplify the analysis, $p_{cl} = 1$ will be used in the following discussion.

Figure 18 displays the acceptance ratio and $\langle \Delta r_M^2 \rangle^{1/2}$ for cluster trial moves as a function of Δ_{cl} for System II with R_{cl} as a parameter from R_M to $2R_M$. The lower limit $R_{cl} = R_M$ implies that no small ions are in the cluster, and thus it corresponds exactly to single-macroion displacements. For a given

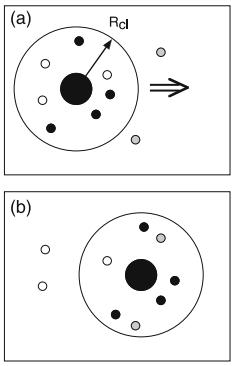

Fig. 17 Schematic illustration of a cluster trial move showing the location of a macroion (*large filled circles*) and counterions (*small circles*) **a** before and **b** after the cluster trial move. Counterions moved together with the macroion are denoted by *filled circles*, those within R_{cl} before the move but not moved by *open circles*, and new ions within R_{cl} after the move by *shaded circles*. In this example, $n_{new} = 7$, $n_{old} = 7$, and $n_{new} - n_{old} = 0$

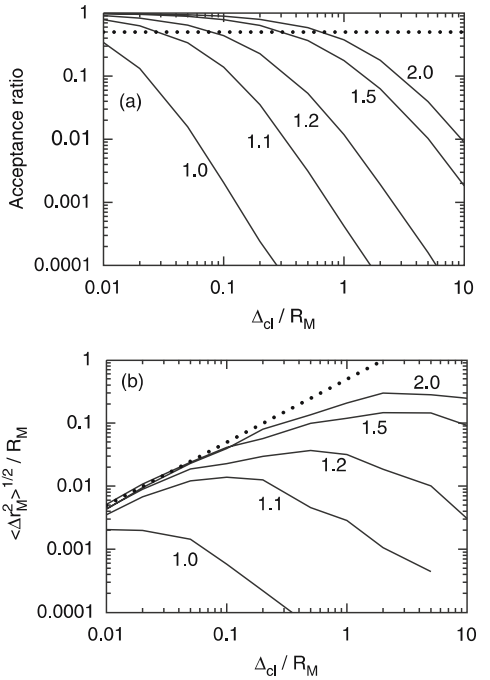

Fig. 18 **a** Macroion acceptance ratio and **b** macroion root-mean-square displacement per MC pass as a function of the cluster displacement parameter for System II from simulations with cluster trial moves using $p_{cl} = 1.0$ and indicated cluster radii R_{cl} in R_M units. At $R_{cl} = R_M$ the single-particle move is recovered. In **a**, the acceptance ratio equal to 0.5, and in **b**, the relation $\langle \Delta r_M^2 \rangle^{1/2} = \Delta_M/2$ are given (*dotted lines*). $N_M = 20$ and $N_{pass} = 10^4$

R_{cl}, both the acceptance ratio and the macroion rms displacement display a dependence on the displacement parameter that is similar to that for single-macroion trial moves. More importantly, however, the acceptance ratio and the macroion rms displacement increase continuously as R_{cl} is increased. The dependence is dramatic when R_{cl} is near R_M and levels off at larger R_{cl}. For example, $\langle \Delta r_M^2 \rangle^{1/2}$ increases 100-fold, corresponding to a 10^4-fold increase in performance, when the cluster displacement is applied with $R_{cl} = 1.5R_M$ at optimal Δ_{cl} ($2–5R_M$) as compared to single-macroion trial moves at its optimal Δ_M ($\approx 0.01R_M$). When defining the simulation efficiency as $\langle \Delta r_M^2 \rangle^{1/2}$ per CPU time, the improvement becomes somewhat smaller due to the larger number of particle pairs to be considered for the many-particle displacement. For example, if Z_r counterions belong to a cluster (an upper limit), the increase in the total CPU time per MC pass is twofold, since the computational load of moving all macroions once increases approximately from $1/Z_r$ to $Z_r(1/Z_r) = 1$ relative to the load of moving all small ions once. Nevertheless, the increase in efficiency is remarkable.

In the case of System III, the improvement is even larger. With the cluster radius $R_{cl} = 1.2R_M$ and $\Delta_{cl} = 5R_M$, the macroion rms displacement becomes $\langle \Delta r_M^2 \rangle^{1/2} \approx R_M$ as compared to $\approx 0.001R_M$ at optimal conditions with the single-macroion trial move (the latter was shown in Fig. 16), resulting in a 10^6-fold increase in performance. By contrast, in the case of Systems I and IV, the macroion acceptance ratio and the rms displacement are only marginally affected by R_{cl} in the interval R_M to $2R_M$.

Features of the cluster displacement approach are further illustrated in Fig. 19, displaying acceptance ratio, macroion rms displacement, and number of cluster members as a function of the electrostatic coupling parameter Γ_{II} at $Z_r = 80$ using the cluster displacement parameter $\Delta_{cl} = R_M$ at different cluster radii R_{cl}. At an electrostatic coupling below $\Gamma_{II} \approx 0.03$, cluster trial moves are slightly inferior to single-macroion trial moves (smaller acceptance ratio and smaller $\langle \Delta r_M^2 \rangle^{1/2}$, see Fig. 19a and b, respectively). As the electrostatic coupling increases with the concomitant accumulation of small ions near the macroion, single-macroion trial moves rapidly deteriorate (see curves labeled "1.0" in Fig. 19a and b). However, by using the cluster displacement procedure and a cluster radius $R_{cl} \approx 1.5R_M$, both the acceptance ratio and $\langle \Delta r_M^2 \rangle^{1/2}$ remain at high levels. There is a noticeable increase in both the acceptance ratio and $\langle \Delta r_M^2 \rangle^{1/2}$ as Γ_{II} is increased at sufficiently large Γ_{II}; hence the cluster displacement procedure becomes *more* efficient as the electrostatic coupling is raised. Consequently, there is an intermediate electrostatic coupling where the performance is least efficient but still highly satisfactory if R_{cl} is selected not too small. Figure 19c shows that the number of cluster members (including the macroion) $N_{cluster}$ increases continuously as Γ_{II} and R_{cl} are increased (with the obvious exception for $R_{cl} = R_M$ when it stays at unity). At $\Gamma_{II} = 0.0222$, $N_{cluster} < 10$, whereas at $\Gamma_{II} = 0.7715$, ca. 70 of the 80 counterions of one macroion are in a shell with thickness $0.2R_M$ at the macroion surface, the latter fact explaining the small sensitivity of $\langle \Delta r_M^2 \rangle^{1/2}$ on R_{cl} appearing for $R_{cl} > 1.2R_M$ at large Γ_{II}.

To conclude, the strong accumulation of counterions near the macroion, being the origin of the failure of single-macroion trial moves, is actually the *key* feature for the success of cluster trial moves. As more counterions are accumulated near a macroion, fewer of them are left in the region between macroions. Hence, even with large cluster displacements, the probability of hard-core overlap is small, facilitating a large total macroion displacement in the configurational space.

Finally, the present comparison has been made for salt-free systems. It should be mentioned that the performance of single-macroion trial moves is essentially salt independent, whereas the performance of cluster trial moves deteriorates upon addition of salt, making the difference in their performances smaller. Nevertheless, the cluster displacement procedure is advantageous up to considerable salt concentrations. A promising method to improve

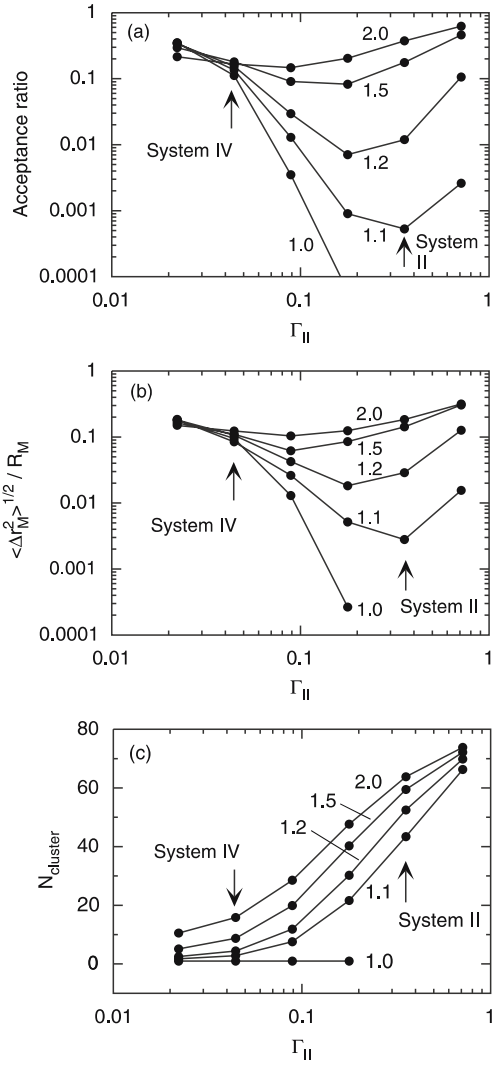

Fig. 19 **a** Macroion acceptance ratio, **b** macroion root-mean-square displacement per MC pass, and **c** number of counterions in a cluster as a function of the electrostatic coupling parameter Γ_{II} at $\Delta_{cl} = R_M$ and indicated radii R_{cl} in R_M units using $Z_r = 80$ and $\phi_M = 0.01$. $\Gamma_{II} = 0.0445$ corresponds to System IV and $\Gamma_{II} = 0.3558$ to System II. $N_M = 20$ and $N_{pass} = 10^4$

such simulations with salt is to invert through the point $\boldsymbol{r}_s + d\boldsymbol{r}/2$ the positions of a macroion and its nearest counterions in a sphere S centered at $\boldsymbol{r}_s$ *and* the positions of the particles in another sphere S' centered at $\boldsymbol{r}_s + d\boldsymbol{r}$ and of the same volume as S. Hence, the contents of spheres S and S' are exchanged (when they are nonoverlapping).

6.3.4
Second-level Cluster Displacement

We have seen that at high electrostatic coupling macroions with nearby counterions (referred to as dressed macroions) form aggregates (Fig. 6a), eventually leading to a phase separation (Fig. 2). The mobility of dressed macroions in such aggregates is very low, even with optimized cluster trial moves. Besides hard-core overlap with nearby dressed macroions, it is energetically unfavorable to displace a macroion M away from another macroion M', with which macroion M shares dressed counterions.

At such conditions, the total macroion displacement may be improved by using a second-level cluster displacement procedure. In this approach, a simultaneous trial displacement of all dressed macroions forming an aggregate (Fig. 20) is performed. Again it is convenient to use a distance criterion for determining the macroions belonging to the same aggregate. In this approach, a mixture of trial moves involving (i) aggregates composed of dressed macroions, (ii) individual dressed macroions, and (iii) individual small ions should be employed.

The performance of joint displacement of several macroions with their counterion dresses has not yet been investigated systematically. However, it has been used to examine the location of the gas–liquid binodal curve of asymmetric electrolyte solutions [22]. Figure 21 shows three snapshots from

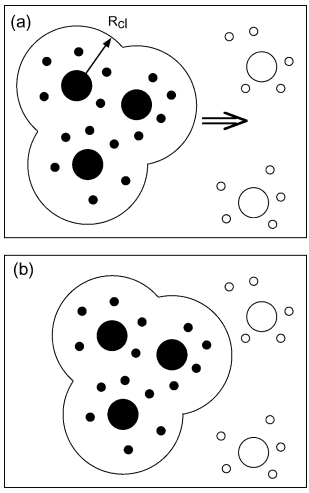

Fig. 20 Schematic illustration of a cluster move of the second level involving a joint displacement of an aggregate composed of macroions and the small ions near these macroions **a** before and **b** after the cluster trial move. Here, three macroions (*large filled circles*) and the small ions (*small filled circles*) within the distance R_{cl} from the macroions form an aggregate subjected to the trial move

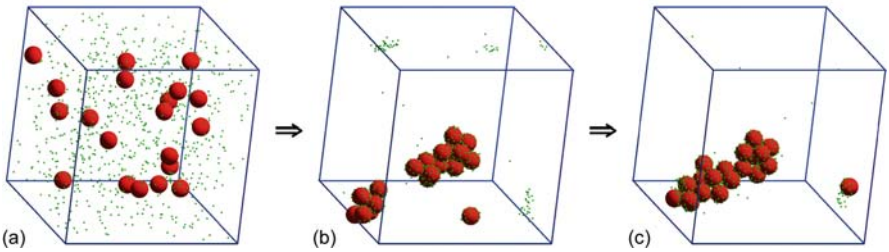

Fig. 21 Snapshots showing the equilibration from an initially homogeneous and unstable configuration where the second level of cluster displacement is used. $N_M = 20$ and the snapshots are separated by $N_{pass} = 5 \times 10^5$ passes. $Z_r = 40$, $\phi_M = 0.01$, and $\Gamma_{II} = 1.940$

the equilibration of an initially homogeneous but unstable solution. After 5×10^5 passes, the initial solution forms two aggregates and one nonaggregated macroion, and then one single aggregate is found. Hence, the final state suggests that a large aggregate is in equilibrium with a very dilute solution, corresponding to a gas–liquid phase separation in the thermodynamic limit. The approach to the equilibrium state is strongly enhanced by allowing displacements of dimers, trimers, etc. of dressed macroions as single entities. With the use of a geometrical criterion and $p_{cl} = 1$, the actual merging/division of aggregates is achieved by cluster trial moves involving one macroion and counterions only.

Acknowledgements Daniel Angelescu, Bo Jönsson, Roland Kjellander, Vladimir Lobaskin, Alexander Lyubartsev, Jurij Rescic, and Håkan Wennerström are gratefully acknowledged for collaboration and fruitful discussions over the years. Special thanks to Samuel Edgecombe, Christian Holm, Bo Jönsson, Vladimir Lobaskin, Alberto Pais, and Khawla Qamhieh for valuable scientific/linguistic comments on the manuscript.

References

1. Evans DF, Wennerström H (1994) The colloidal domain where physics, chemistry, biology, and technology meet. VCH, New York
2. Arora AK, Tata BVR (1996) In: Rajagopalan R (ed) Complex fluids and fluid microstructures. VCH, New York
3. Derjaguin BV, Landau L (1941) Acta Physiochim (USSR) 14:633
4. Verwey EJ, Overbeek JTG (1948) Theory of the stability of lyophobic colloids. Elsevier, Amsterdam
5. Khan A, Fontell K, Lindman B (1984) J Colloid Interface Sci 101:193
6. Kang C, Khan A (1993) J Colloid Interface Sci 156:218
7. Tata BVR, Rajalakshmi M, Arora AK (1992) Phys Rev Lett 69:3778
8. Ito K, Yoshida H, Ise H (1994) Science 263:66
9. Larsen AE, Grier DG (1997) Nature 385:230
10. Grier DG (1998) Nature 393:621

11. Vlachy V (1999) Annu Rev Phys Chem 50:145
12. Hansen J-P, Löwen H (2000) Annu Rev Phys Chem 51:209
13. Belloni L (2000) J Phys Condens Matter 12:R549
14. Bhuiyan LB, Vlachy V, Outhwaite CW (2002) Int Rev Phys Chem 21:1
15. Allahyarov E, D'Amico I, Löwen H (1998) Phys Rev Lett 81:1334
16. Grønbech-Jensen N, Beardmore KM, Pincus P (1998) Physica A 261:74
17. Wu J, Bratko D, Prausnitz JM (1998) Proc Natl Acad Sci 95:15169
18. Linse P, Lobaskin V (1999) Phys Rev Lett 83:4208
19. Wu J, Bratko D, Blanch HW, Prausnitz JM (1999) J Chem Phys 111:7084
20. Linse P, Lobaskin V (2000) J Chem Phys 112:3917
21. Hribar B, Vlachy V (2000) Biophys J 78:694
22. Linse P (2000) J Chem Phys 113:4359
23. Messina R, Holm C, Kremer K (2000) Phys Rev Lett 85:872
24. Messina R, Holm C, Kremer K (2000) Europhys Lett 51:461
25. Belloni L (2002) J Phys Condens Matter 14:9323
26. Linse P (2002) J Phys Condens Matter 14:13449
27. Lobaskin V, Qamhieh K (2003) J Phys Chem B 107:8022
28. Angelescu DG, Linse P (2003) Langmuir 19:9661**?**
29. Kjellander R, Marcelja S (1984) Chem Phys Lett 112:49
30. Outhwaite CW, Molero M (1992) Chem Phys Lett 197:643
31. Rouzina I, Bloomfield VA (1996) J Phys Chem 100:9977
32. van Roij R, Hansen J-P (1997) Phys Rev Lett 79:3082
33. Tokuyama M (1999) Phys Rev E 59:R2550
34. Levin Y (1999) Physica A 265:432
35. van Roij R, Dijkstra M, Hansen J-P (1999) Phys Rev E 59:2010
36. Warren PB (2000) J Chem Phys 112:4683
37. Naji A, Netz RR (2004) Eur Phys J E 13:43
38. Guldbrand L, Jönsson B, Wennerström H, Linse P (1984) J Chem Phys 80:2221
39. Moreira AG, Netz RR (2001) Phys Rev Lett 87:078301
40. Moreira AG, Netz RR (2002) Eur Phys J E 8:33
41. Guldbrand L, Nilsson LG, Nordenskiöld L (1986) J Chem Phys 85:6686
42. Grønbech-Jensen N, Mashl RJ, Bruinsma RF, Gelbart WM (1997) Phys Rev Lett 78:2477
43. Ha B-Y, Liu AJ (1997) Phys Rev Lett 79:1289
44. Ha B-Y, Liu AJ (1998) Phys Rev Lett 81:1011
45. Podgornik R, Parsegian VA (1998) Phys Rev Lett 80:1560
46. Shklovskii BI (1999) Phys Rev Lett 82:3268
47. Solis FJ, Olverade la Cruz M (1999) Phys Rev E 60:4496
48. Deserno M, Arnold A, Holm C (2003) Macromolecules 2003:249
49. Stevens MJ (1999) Phys Rev Lett 82:101
50. Stevens MJ, Kremer K (1995) J Chem Phys 103:1669
51. Winkler RG, Gold M, Reineker P (1998) Phys Rev Lett 80:3731
52. Khan MO, Jönsson B (1999) Biopolymers 49:121
53. Pais AACC, Miguel MG, Linse P, Lindman B (2002) J Chem Phys 117:1385
54. Orkoulas G, Kumar SK, Panagiotopoulos AZ (2003) Phys Rev Lett 90:048303
55. Stevens MJ (2001) Biophys J 80:130
56. Yan Q, de Pablo JJ (2003) Phys Rev Lett 67:061803
57. Schneider S, Linse P (2004) Macromolecules 37:3850
58. Linse P (1990) J Chem Phys 93:1376
59. Ewald P (1921) Ann Phys 64:253

60. Gordon HL, Valleau JP (1995) Mol Simul 14:361
61. Linse P, Lobaskin V (2000) In: Engquist B, Johnsson L, Hammill M, Short F (eds) Simulation and Visualization on the Grid. Lecture Notes in Computational Science and Engineering, vol 13. Springer, Berlin Heidelberg New York, p 165
62. McMillan WG, Mayer JE (1945) J Chem Phys 13:276
63. Hill TL (1960 and 1986) An Introduction to Statistical Thermodynamics. Dover, New York
64. Allen MP, Tildelsley DJ (1987) Computer simulation of liquids. Oxford, New York
65. Frenkel D, Smit B (1996) Understanding Molecular Simulation: From Algorithms to Applications. Academic, San Diego
66. Sokal AD (1996) Lectures at the Cargèse Summer School on "Functional Integration: Basis and Applications"
67. Mille M (1977) J Colloid Interface Sci 59:211
68. Frahm J, Diekmann S (1979) J Colloid Interface Sci 70:440
69. Bentz J (1981) J Colloid Interface Sci 80:179
70. Linse P, Gunnarsson G, Jönsson B (1982) J Phys Chem 86:413
71. Lampert MA, Martinelli RU (1985) Chem Phys Lett 121:121
72. Amaro ML, Bhuiyan LB, Outhwaite CW (1995) Mol Phys 86:725
73. Degreve L, Lozada-Cassou M (1995) Mol Phys 86:758
74. Rebolj N, Kristl J, Kalyuzhnyi YV, Vlachy V (1997) Langmuir 13:3646
75. Bell GM (1964) J Trans Faraday Soc 60:1752
76. Bell GM, Dunning AJ (1970) J Trans Faraday Soc 66:500
77. Bratko D, Lindman B (1985) J Phys Chem 89:1437
78. Jönsson B, Wennerström H, Nilsson PG, Linse P (1986) Colloid Polym Sci 264:77
79. Gunnarsson G, Jönsson B, Wennerström H (1980) J Phys Chem 84:3114
80. Jönsson B, Wennerström H (1981) J Colloid Interface Sci 80:482
81. Wennerström H, Jönsson B, Linse P (1982) J Chem Phys 76:4665
82. Marcus RA (1955) J Chem Phys 23:1057
83. da Silva FLB, Bogren D, Söderman O, Åkesson T, Jönsson B (2002) J Phys Chem B 106:3515
84. Messina R, Holm C, Kremer K (2001) Eur Phys J D 4:363
85. Messina R (2002) Physica A 308:57
86. Linse S, Jönsson B, Chazin WJ (1995) Proc Natl Acad Sci USA 92:4748
87. Linse P (1986) J Phys Chem 90:6821
88. Messina R (2002) J Chem Phys 117:11062
89. von Grünberg HH, Belloni L (2000) Phys Rev E 62:2493
90. Wallin T, Linse P (1996) J Phys Chem 100:17873
91. Wallin T, Linse P (1996) Langmuir 12:305
92. Wallin T, Linse P (1997) J Phys Chem B 101:5506
93. Mateescu EM, Jeppesen C, Pincus P (1999) Europhys Lett 46:493
94. Nguyen TT, Shklovskii BI (2001) Physica A 293:324
95. Akinchina A, Linse P (2002) Macromolecules 35:5183
96. Akinchina A, Linse P (2003) Macromolecules 107:8011
97. Woodward C, Jönsson B, Åkesson T (1988) J Chem Phys 89:5145
98. Granfeldt MK, Jönsson B, Woodward CE (1991) J Phys Chem 95:4819
99. Striolo A, Bratko D, Wu JZ, Elvassore N, Blanch HW, Prausnitz JM (2002) J Chem Phys 116:7733
100. Allahyarov E, Löwen H, Louis AA, Hansen JP (2002) Europhys Lett 57:731
101. Allahyarov E, Löwen H, Hansen JP, Louis AA (2003) Phys Rev E 67:051404
102. Lund M, Jönsson B (2003) Biophys J 85:2940

103. Linse P, Jönsson B (1983) J Chem Phys 78:3167
104. Vlachy V, Marshall CH, Haymet ADJ (1989) J Am Chem Soc 111:4160
105. Hribar B, Kalyuzhnyi YV, Vlachy V (1996) Mol Phys 87:1317
106. Hribar B, Vlachy V (1997) J Phys Chem B 101:3457
107. Hribar B, Krienke H, Kalyuzhnyi YV, Vlachy V (1997) J Mol Liq 73–74:277
108. Lobaskin V, Linse P (1998) J Chem Phys 109:3530
109. Lobaskin V, Linse P (1999) J Chem Phys 111:4300
110. Neu JC (1999) Phys Rev E 82:1072
111. Sader JE, Chan DYC (1999) J Colloid Interface Sci 203:268
112. Rescic J, Linse P (2000) J Phys Chem B 104:7852
113. Skepö M, Linse P (2002) J Phys Chem B 36:508
114. Carlsson F, Malmsten M, Linse P (2003) J Am Chem Soc 125:3140
115. Jardat M, Cartailler T, Turq P (2001) J Chem Phys 115:1066
116. Landau LD, Lifshitz EM (1960) Electrodynamics of continuous media. Pergamon, Bristol
117. Russel WB, Saville DA, Schowalter WR (1989) Colloidal dispersions. Cambridge University Press, New York
118. Adams DJ (1979) Chem Phys Lett 62:329
119. Ceperly D (1980) NRCC Proc no 9. National Technical Information Service, Springfield, VA
120. de Leeuw SW, Perram JW, Smith ER (1980) Proc R Soc Lond A 373:27
121. Romero-Enrique JM, Orkoulas G, Panagiotopoulos AZ, Fisher ME (2000) Phys Rev Lett 85:4558
122. de Leeuw SW, Perram JW, Smith ER (1980) Proc R Soc Lond A 373:57
123. Perram JW, Petersen HG, de Leeuw SW (1988) Mol Phys 65:875
124. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089
125. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577
126. Petersen HG (1995) J Chem Phys 103:3668
127. Deserno M, Holm C (1998) J Chem Phys 109:7678
128. Kolafa J, Perram JW (1992) Molec Sim 9:351
129. Linse P (1999) J Chem Phys 110:3493
130. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) J Chem Phys 21:1087

Adv Polym Sci (2005) 185: 163–225
DOI 10.1007/b136796

Published online: 2 September 2005

Simple Dipolar Fluids as Generic Models for Soft Matter

J.-J. Weis (✉) · D. Levesque

Laboratoire de Physique Théorique, Université Paris-Sud, Bâtiment 210, UMR 8627, 91405 Orsay Cedex, France
jean-jacques.weis@th.u-psud.fr, dominique.levesque@th.u-psud.fr

Abstract The physical properties, based on simulation results, of model fluids and solids bearing an electric or magnetic point dipole moment are described. Comparison is made with experimental data on ferrofluids and electro- or magneto-rheological fluids. The qualitative agreement between experiment and simulation shows the interest of these simple models for the comprehension of physical systems where the dipolar interaction dominates.

Keywords Dipolar fluids · Computer simulations · Ewald sums · Electrorheology

Abbreviations

CMM	cell multipole method
DHS	dipolar hard spheres
DSS	dipolar soft spheres
ER	electrorheological
FFT	fast Fourier transform

FMM fast multipole method
GB Gay–Berne
HS hard spheres
LC liquid crystal
LEBC Lees-Edwards boundary conditions
LJ Lennard-Jones
MC Monte Carlo
MD molecular dynamics
MMM M body method
MR magnetorheological
PME particle-mesh Ewald
P3M particle-particle particle-mesh
Q2D quasi two-dimensional
SPME smooth particle-mesh Ewald
TEM transmission electron microscopy
b.c. boundary conditions
l-g liquid-gas

1 Introduction

The systems of dipolar particles considered in this review are simple models for fluids or solids whose molecules bear a permanent dipole moment. They can also serve as credible models for suspensions like ferrofluids [1] or electro(magneto)rheological fluids [2, 3]. For spherically symmetric dipolar particles the interaction between a pair of particles separated by a distance r_{ij} and bearing dipole moments $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ is given by

$$v(\boldsymbol{r}_{ij}, \boldsymbol{\mu}_i, \boldsymbol{\mu}_j) = v_{\mathrm{sr}}(r_{ij}) + \frac{1}{r_{ij}^p}\left[\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j - \frac{p(\boldsymbol{\mu}_i \cdot \boldsymbol{r}_{ij})(\boldsymbol{\mu}_j \cdot \boldsymbol{r}_{ij})}{\boldsymbol{r}_{ij}^2}\right] \tag{1}$$

where $v_{\mathrm{sr}}(r)$ is an isotropic short-range potential mostly taken to be a hard sphere potential of diameter σ, a soft sphere potential $4\varepsilon(\sigma/r)^{12}$ or a Lennard-Jones potential $4\varepsilon[(\sigma/r)^{12} - (\sigma/r)^6]$; $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$ is the vector joining the centres of mass of the particles. The cases $p = 3$ and $p = 2$ correspond to dipolar interactions satisfying three-dimensional (3D) and two-dimensional (2D) electrostatics, respectively. Two-dimensional systems where the particle centres lie in a plane but the dipole moments can rotate in full 3D space will be referred to as Q2D systems. For molecular systems, as for instance liquid crystals, the short range potential v_{sr} will also depend on the relative orientations of the two molecules. The study of these systems has been developed to determine the diversity of phase diagrams generated by the sole association of excluded volume effects and dipolar interactions. While a gas phase is always present at high temperature, different phases, liquid, solid, liquid-

crystal, ferroelectric or polymer-like, can occur at low temperature or high density. The standard liquid state theories [4] can account quantitatively for the high temperature fluid phases, but only numerical simulations are capable of giving reliable results for the appearance and nature of the phases which exist when the temperature is lowered or the density increased. In view of the vast amount of literature devoted to the properties of dipolar fluids and solids and their suspensions, we shall limit this review to simulation results, mainly obtained after 1995, focusing on qualitative aspects. For reasons of space limitations no comparison with theoretical developments is included but, when possible, experimental observations are mentioned to emphasize the motivation of the simulations and the effects searched for, as well as to judge the relevance of the models used. Complementary reviews of simple dipolar systems are given in [5, 6] (also contains theoretical aspects) and [7] (Q2D dipolar systems) and for ER fluids in [8, 9] (ER fluids prior to 1996).

Throughout the review we will use reduced units which differ according to whether the short range potential is strictly of a hard core or soft core type with or without dispersion forces (soft core or LJ interaction). For hard core potentials a thermodynamic state is characterized by the reduced density $\rho^* = N\sigma^3/V$ (N number of particles, V volume, σ characteristic length of the hard core) and a reduced dipole moment $\mu^* = (\mu^2/\sigma^3 kT)^{1/2}$ (μ dipole moment, T temperature, k Boltzmann constant), or a reduced temperature $T^* = 1/\mu^{*2}$. For soft core potentials, a state is defined by the reduced density $\rho^* = N\sigma^3/V$, the reduced temperature $T^* = kT/\varepsilon$ and the reduced dipole moment $\mu^* = (\mu^2/\sigma^3\varepsilon)^{1/2}$, where σ and ε are the parameters defining the range and the strength of the short range potential. It is also convenient to define the dimensionless coupling parameter $\lambda = \mu^{*2}$ or the related parameter $\bar{\lambda} = 2\lambda$, the ratio of the energy of two aligned dipoles at close contact to the thermal energy.

When dealing with simple dipolar systems we will take $4\pi\varepsilon_0 = 1$ (ε_0 permittivity of vacuum); SI units will be used in Sect. 5 on ER or MR fluids.

2 Methods

The standard way to overcome the surface effects which result from the finite size of systems studied in numerical simulations is the use of periodic boundary conditions. With this requirement, systems of 1000-10 000 particles interacting by potentials having a range of the order of a few particle diameters are sufficient to perform simulations where the bias on the computation data, induced by the finite size effects is at an acceptable level of $\sim 1\%$. However, when the interactions between the particles are of Coulombic and/or dipolar types, their long range cannot be neglected because it is at the origin

of their main specific physical properties, screening and dielectric phenomena.

There are two methods, which give equivalent results, for making the periodic boundary conditions compatible with the conservation of the physical properties of Coulombic and dielectric systems. In one of these methods, the potential between charges is the potential obtained by solving the Poisson equation for a point charge in a volume, with periodic boundary conditions, identical to that considered in the simulations. However, since in a finite volume, a solution of the Poisson equation exists only for a neutral system, it is convenient to associate with the point charge a uniform neutralizing background filling the system volume. This background contributes explicitly to the energy of a system of N charges q_j ($j = 1, ..., N$) only when $\sum_j q_j \neq 0$. Thus, the potential at point $\boldsymbol{r}$ of a parallelepipedic simulation cell of volume V, defined by the basis vectors $\boldsymbol{t}_1$, $\boldsymbol{t}_2$, and $\boldsymbol{t}_3$, containing a neutral system of N charges is given by [10]

$$\phi(\boldsymbol{r}) = \frac{4\pi}{V} \sum_{\boldsymbol{G}\neq 0} \frac{1}{G^2} \sum_{j=1}^{N} q_j \exp\left(\mathrm{i}\boldsymbol{G}\cdot(\boldsymbol{r}-\boldsymbol{r}_j)\right). \tag{2}$$

where $\boldsymbol{r}_j$ is the position of the charge j. The reciprocal space vectors $\boldsymbol{G}$ are equal to $2\pi\,({}^t\boldsymbol{t})^{-1}\cdot\boldsymbol{n}$, $\boldsymbol{n}$ being a vector with integer components n_x, n_y and n_z ranging from $-\infty$ to $+\infty$ and $({}^t\boldsymbol{t})^{-1}$ is the inverse of the transpose of the 3×3 matrix $\boldsymbol{t}$, the 3 columns of which are equal to $\boldsymbol{t}_1$, $\boldsymbol{t}_2$, and $\boldsymbol{t}_3$.

In the other method, the space is considered filled with an infinite lattice of replicas of the system charges. Considering that the system is neutral and that the charges and those in the replicas interact by the Coulomb potential, it is easy to write the formal expression of the total Coulomb energy:

$$E_C = \frac{1}{2} \sum_{i,j=1}^{N} {\sum_{\boldsymbol{n}}}' \frac{q_j q_i}{|\boldsymbol{r}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|} \tag{3}$$

where the prime indicates that, when $\boldsymbol{n} = 0$, i must be different from j. The energy of a system of N dipolar particles with dipole moment $\boldsymbol{\mu}_i$ can be derived from the previous expression of E_C for systems with periodic boundary conditions by substituting the charges q_i with the operators $-\boldsymbol{\mu}_i.\nabla_i$

$$E = \frac{1}{2} \sum_{i,j=1}^{N} {\sum_{\boldsymbol{n}}}' \frac{1}{|\boldsymbol{r}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|^3} \left[\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j - 3\,\frac{\boldsymbol{\mu}_i\cdot(\boldsymbol{r}_{ij}+\boldsymbol{t}\cdot\boldsymbol{n})\;\;\boldsymbol{\mu}_j\cdot(\boldsymbol{r}_{ij}+\boldsymbol{t}\cdot\boldsymbol{n})}{|\boldsymbol{r}_{ij}+\boldsymbol{t}\cdot\boldsymbol{n}|^2} \right] \tag{4}$$

As has been extensively discussed in the literature, these expressions are conditionally convergent series and appropriate procedures have been developed to guarantee their physically correct evaluations [11–15]. Different, though

equivalent, expressions of the Coulomb or dipolar energy can result from these procedures, the well known Ewald sum [11] and the so-called Lekner sum [12, 16]. The evaluation of E can be made by either performing the Ewald or Lekner sums directly in Eq. 4 or by first making the sums in the Coulomb energy (Eq. 3) and applying the operators $-\boldsymbol{\mu}_i.\nabla_i$.

In the following we will give a summary of the energy expressions obtained by the Ewald and Lekner summation methods, for 3D and 2D point dipolar systems periodic in all spatial directions or periodic in only some directions and finite in the others. The energies obtained with the Lekner summation are often quite lengthy and will be omitted when available in the original papers.

3D Systems

Periodic Boundary Conditions

The Ewald expression for the energy of a fully periodic 3D system of N dipolar particles with dipole moments $\boldsymbol{\mu}_i$ is given by [11]

$$
\begin{aligned}
E = & \frac{1}{2} \sum_{i,j=1}^{N} {\sum_{\boldsymbol{n}}}' \frac{1}{|\boldsymbol{r}_{ij} + \boldsymbol{t} \cdot \boldsymbol{n}|^3} \Bigg[\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j \, B(\alpha |\boldsymbol{r}_{ij} + \boldsymbol{t} \cdot \boldsymbol{n}|) \\
& - 3 \frac{\boldsymbol{\mu}_i \cdot (\boldsymbol{r}_{ij} + \boldsymbol{t} \cdot \boldsymbol{n}) \quad \boldsymbol{\mu}_j \cdot (\boldsymbol{r}_{ij} + \boldsymbol{t} \cdot \boldsymbol{n})}{|\boldsymbol{r}_{ij} + \boldsymbol{t} \cdot \boldsymbol{n}|^2} \, C(\alpha |\boldsymbol{r}_{ij} + \boldsymbol{t} \cdot \boldsymbol{n}|) \Bigg] \\
& + \frac{2\pi}{V} \sum_{i,j=1}^{N} \sum_{\boldsymbol{G} \neq 0} \mathrm{e}^{-G^2/(4\alpha^2)} \frac{(\boldsymbol{\mu}_i \cdot \boldsymbol{G}) \, (\boldsymbol{\mu}_j \cdot \boldsymbol{G})}{G^2} \, \exp(\mathrm{i} \boldsymbol{G} \cdot \boldsymbol{r}_{ij}) \\
& - \frac{2\alpha^3}{3\sqrt{\pi}} \sum_{i=1, \cdots, N} \boldsymbol{\mu}_j^2 \, ,
\end{aligned}
\tag{5}
$$

where α is the parameter regulating the relative convergence of the sums in direct r- and reciprocal k-space and $B(r)$ and $C(r)$ are defined by

$$
B(r) = \mathrm{erfc}(\alpha r) + \frac{2\alpha r}{\sqrt{\pi}} \exp(-\alpha^2 r^2) \tag{6}
$$

and

$$
C(r) = \mathrm{erfc}(\alpha r) + \frac{2\alpha r}{\sqrt{\pi}} \exp(-\alpha^2 r^2) \left(1 + \frac{2}{3} \alpha^2 r^2 \right) \tag{7}
$$

with $\mathrm{erfc}(x)$ the complementary error function.

The Lekner expression for the energy for a 3D Coulomb periodic system has been computed in [17]; that for a 3D dipolar periodic system in an ortho-

rhombic simulation cell follows

$$E = \frac{1}{2}\sum_{i\neq j=1}^{N} R(x_{ij}, y_{ij}, z_{ij}) + S(x_{ij}, y_{ij}, z_{ij}) + T(x_{ij}, y_{ij}, z_{ij}) \qquad (8)$$
$$+ U(x_{ij}, y_{ij}, z_{ij}) + V(y_{ij}, z_{ij}) + W(y_{ij}, z_{ij}) + E_{\text{self}}$$

where E_{self} is the constant contribution of the dipole self-energy.

The different functions $R(x_{ij}, y_{ij}, z_{ij})$, $S(y_{ij}, z_{ij})$ ... are given by

$$R(x_{ij}, y_{ij}, z_{ij}) = \frac{4\pi}{L_x^3}\sum_{n_x\neq 0, n_y, n_z} (\mu_i^y \mu_j^y + \mu_i^z \mu_j^z) \mathrm{e}^{\mathrm{i}2\pi\, n_x x_{ij}/L_x} \qquad (9)$$
$$\times \frac{|n_x|}{|\boldsymbol{\tau}_{ij}^{n_y n_z}|} K_1\left(2\pi |n_x||\boldsymbol{\tau}_{ij}^{n_y n_z}|\right)$$

$$S(x_{ij}, y_{ij}, z_{ij}) = -\frac{8\pi^2}{L_x^5}\sum_{n_x\neq 0, n_y, n_z} n_x^2 (\mu_i^y (y_{ij} + n_y L_y) + \mu_i^z (z_{ij} + n_z L_z))$$
$$\times (\mu_j^y (y_{ij} + n_y L_y) + \mu_j^z (z_{ij} + n_z L_z)) K_2(2\pi |n_x||\boldsymbol{\tau}_{ij}^{n_y n_z}|) \frac{\mathrm{e}^{\mathrm{i}2\pi n_x x_{ij}/L_x}}{\boldsymbol{\tau}_{ij}^{n_y n_z\, 2}}$$

$$T(x_{ij}, y_{ij}, z_{ij}) = \frac{8\pi^2}{L_x^3}\sum_{n_x\neq 0, n_y, n_z} \mu_i^x \mu_j^x n_x^2 \mathrm{e}^{\mathrm{i}2\pi n_x x_{ij}/L_x}$$
$$\times K_0(2\pi |n_x||\boldsymbol{\tau}_{ij}^{n_y n_z}|)$$

$$U(x_{ij}, y_{ij}, z_{ij}) = \frac{16\pi^2}{L_x^4}\sum_{n_x>0, n_y, n_z} \sin(2\pi n_x x_{ij}/L_x)$$
$$\times \left\{(\mu_i^x \mu_j^y + \mu_i^y \mu_j^x)(y_{ij} + n_y L_y) + (\mu_i^x \mu_j^z + \mu_i^z \mu_j^x)(z_{ij} + n_z L_z)\right\}$$
$$\times \frac{n_x^2}{|\boldsymbol{\tau}_{ij}^{n_y n_z}|} K_1(2\pi |n_x||\boldsymbol{\tau}_{ij}^{n_y n_z}|)$$

$$V(y_{ij}, z_{ij}) = \frac{8\pi^2}{L_x L_y^3}\sum_{n_y>0, n_z} (\mu_i^y \mu_j^z + \mu_i^z \mu_j^y)(z_{ij} + n_z L_z)$$
$$\times \sin(2\pi n_y y_{ij}/L_y) \frac{n_y L_y}{|(z_{ij} + n_z L_z)|} \mathrm{e}^{-2\pi n_y |z_{ij} + n_z L_z|/L_y}$$

$$W(y_{ij}, z_{ij}) = \frac{4\pi^2}{L_x L_y^2}\sum_{n_y, n_z} (\mu_i^y \mu_j^y - \mu_i^z \mu_j^z) |n_y|\, \mathrm{e}^{\mathrm{i}2\pi n_y y_{ij}/L_y}$$
$$\times \mathrm{e}^{-2\pi |n_y||z_{ij} + n_z L_z|/L_y}$$

where K_ν are the modified Bessel functions of second type and $\boldsymbol{\tau}_{ij}^{n_y n_z}$ is a 2D vector of the yz plane with components $(y_{ij} + n_y L_y)/L_x$ and $(z_{ij} + n_z L_z)/L_x$.

The sum on n_y can be performed explicitly in $W(y_{ij}, z_{ij})$ with the help of the relation

$$\sum_{n_y} e^{i2\pi n_y y_{ij}/L_y} e^{-2\pi|n_y||z_{ij}+n_z L_z|/L_y} = \frac{\sinh(2\pi|z_{ij}+n_z L_z|/L_y)}{\cosh(2\pi|z_{ij}+n_z L_z|/L_y) - \cos(2\pi y_{ij}/L_y)} . \tag{10}$$

The preceding Ewald energy expressions have been written in a form corresponding to the so-called tin foil boundary condition which supposes, when the summation of the pair interactions is performed, that, on a macroscopic scale, the system and its replicas are surrounded by a conducting medium. If the latter medium is a dielectric with a dielectric constant ε', it will be polarized by the total dipole moment $\boldsymbol{M} = \sum \boldsymbol{\mu}_i$ of the dipolar particles and a novel shape dependent contribution E_P resulting from this polarization effect should be added to E [18]:

$$E_P = \frac{2\pi}{(2\varepsilon' + 1)V} \boldsymbol{M}^2 . \tag{11}$$

Hypersphere

The use of periodic boundary conditions is not the only way to remove surface effects in finite charged or dipolar systems. An alternative is to use for the simulation volume the 3D surface S_3 of a four-dimensional (4D) sphere (hypersphere) which has no boundary [19–21]. The specific physical properties of the charged systems are preserved if the charges interact by a potential solution of the Poisson equation which for S_3 is

$$\Delta_{S_3} V(M, M_i) = -4\pi q_i \left\{ \delta_{S_3}(M, M_i) - \frac{1}{2\pi^2 R^3} \right\} \tag{12}$$

where Δ_{S_3} denotes the Laplacian operator and δ_{S_3} the Dirac distribution in S_3. The origin of the coordinates is taken at the centre of the 4D sphere. M and M_i are the locations of any point and the charge q_i in S_3 and R is the radius of the 4D sphere. Similarly to the case of the finite 3D system, a term representing a charged uniform background of total charge opposite to that of the point charge has been introduced in order to maintain neutrality. The main advantage of considering charge systems in S_3 is the fact that the solution of the Poisson equation has a closed analytical form, leading to a potential $\phi_{S_3}(\boldsymbol{X}_i, \boldsymbol{X}_j)$ between two charges q_i and q_j located at $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ ($\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ are 4D vectors specifying the positions M_i and M_j of the charges), respectively, given by [19]

$$\phi_{S_3}(\boldsymbol{X}_i, \boldsymbol{X}_j) = \phi_{S_3}(\psi_{ij}) = \frac{q_i q_j}{\pi R} \left\{ (\pi - \psi_{ij}) \cot \psi_{ij} - \frac{1}{2} \right\} \tag{13}$$

where $\psi_{ij} = \operatorname{acos}(\boldsymbol{X}_i \cdot \boldsymbol{X}_j / R^2)$.

Substituting $-\boldsymbol{\mu}_i \cdot \nabla_{\boldsymbol{X}_i}$ for q_i, where $\boldsymbol{\mu}_i$ is a 4D vector tangent to the surface of the 4D hypersphere at $\boldsymbol{X}_i$ having a modulus equal to the dipole moment μ_i of particle i, the dipole-dipole potential gives [20–22]

$$E = \frac{1}{2}\sum_{i\neq j}^{N} \phi_{S_3,dd}(\boldsymbol{X}_i, \boldsymbol{X}_j) = \frac{1}{2}\sum_{i\neq j}^{N} \phi_{S_3,dd}(\psi_{ij}) \tag{14}$$

$$= \frac{1}{2}\sum_{i\neq j}^{N} \frac{1}{\pi R^3}\left\{ 2\,\frac{(\boldsymbol{\mu}_i \cdot \boldsymbol{X}_j)\,(\boldsymbol{\mu}_j \cdot \boldsymbol{X}_i)}{R^2\,\sin^2\psi_{ij}} + f(\psi_{ij})\left[\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j + 3\cos\psi_{ij}\frac{(\boldsymbol{\mu}_i \cdot \boldsymbol{X}_j)\,(\boldsymbol{\mu}_j \cdot \boldsymbol{X}_i)}{R^2 \sin^2\psi_{ij}}\right]\right\}.$$

with

$$f(\psi_{ij}) = \frac{1}{\sin\psi_{ij}}\left(\cot\psi_{ij} + \frac{\pi - \psi_{ij}}{\sin^2\psi_{ij}}\right). \tag{15}$$

How to sample uniformly a point M on the surface of a hypersphere and how to make a trial displacement of a dipole is explained in the Appendix of [22]. The orientational order of dipoles can be described by expanding the two-particle pair correlation function on a complete set of rotational invariants in 4D space. The first projections are:

$$\phi^{000}(1,2) = 1 \tag{16}$$

$$\phi^{110}(1,2) = \boldsymbol{s}_1 \odot \boldsymbol{s}_2$$

$$\phi^{112}(1,2) = 3(\boldsymbol{s}_1 \odot \boldsymbol{t}_{12}(1))(\boldsymbol{s}_2 \odot \boldsymbol{t}_{12}(2)) - \boldsymbol{s}_1 \odot \boldsymbol{s}_2$$

where $\boldsymbol{s}_1 = \boldsymbol{\mu}_1/|\boldsymbol{\mu}_1|$, $\odot$ denotes the scalar product in S_3 and $\boldsymbol{t}_{12}(1)$ is a unit vector tangent to the geodesic $M_1\,M_2$ at M_1 and oriented from M_1 towards M_2.

In terms of the scalar products in 4D space, Eqs. 16 can be expressed as [20]

$$\phi^{110}(1,2) = \boldsymbol{s}_1 \cdot \boldsymbol{s}_2 - \frac{1}{1+\cos\psi_{12}}\,\frac{\boldsymbol{s}_1 \cdot \boldsymbol{X}_2}{R}\,\frac{\boldsymbol{s}_2 \cdot \boldsymbol{X}_1}{R} \tag{17}$$

and

$$\phi^{112}(1,2) = -\,\boldsymbol{s}_1 \cdot \boldsymbol{s}_2 - \frac{2+\cos\psi_{12}}{\sin^2\psi_{12}}\,\frac{\boldsymbol{s}_1 \cdot \boldsymbol{X}_2}{R}\,\frac{\boldsymbol{s}_2 \cdot \boldsymbol{X}_1}{R}. \tag{18}$$

The knowledge of ϕ^{110} and ϕ^{112} are sufficient to determine the thermodynamic and dielectric properties of the system. Thus, the energy of a pair of particles is given by [20]

$$\phi_{S_3,dd}(\psi_{12}) = \frac{\mu_1\mu_2}{3\pi R^3}\big[f(\psi_{12})\big(2\phi^{110}(1,2) - \phi^{112}(1,2)\big) - 2(1+\cos\psi_{12})\big(\phi^{110}(1,2) + \phi^{112}(1,2)\big)\big] \tag{19}$$

The dielectric constant corresponding to the dipole-dipole internal energy (Eq. 14) is related to fluctuations of the total dipole moment of the system $\boldsymbol{M} = \sum_i \boldsymbol{\mu}_i$ by

$$\frac{4\pi}{3V}\left\langle \boldsymbol{M}^2 \right\rangle = \frac{\varepsilon - 1}{\varepsilon}. \tag{20}$$

Unfortunately, Eq. 20 is of limited practical use since, at high dipole moments, the value of ε depends sensitively on the statistical error on $\langle \boldsymbol{M}^2 \rangle$. Other expressions for the dielectric constant can be derived by considering an internal energy of the form [20]

$$E^{(\lambda)} = E + 2\,\lambda(\lambda - 2)\frac{4\pi}{3V}\left\langle \boldsymbol{M}^2 \right\rangle \tag{21}$$

where E is given by Eq. 14. The additional term does not modify the structural and thermodynamic properties of the system in the thermodynamic limit but affects the fluctuations of the total dipole moment yielding for the expression of the dielectric constant

$$\begin{aligned} \frac{4\pi}{3V}\left\langle \boldsymbol{M}^2 \right\rangle &= \frac{\varepsilon - 1}{1 + (\lambda - 1)^2(\varepsilon - 1)} \\ &= \frac{(\varepsilon - 1)(2\varepsilon' + 1)}{2\varepsilon' + \varepsilon} \end{aligned} \tag{22}$$

where the second expression is obtained from the substitution $(\lambda - 1)^2 = 1/(2\varepsilon' + 1)$. In this form the relation between ε and $\langle \boldsymbol{M}^2 \rangle$ is formally identical to that of a periodic system with a dielectric medium ε' surrounding a large sphere representing a periodic replica of a 3D simulation cell. In [23] Caillol notes that strong size effects affect the thermodynamic, structural and dielectric properties with the choice $\varepsilon' \neq 0$ and rather recommends the use of $\varepsilon' = 0$ ($\lambda = 0$). The drawback of Eq. 20 can be circumvented by relating the dielectric constant to the dipole moment fluctuation $\langle \boldsymbol{m}^2(\psi) \rangle$ in a subvolume of the system [23]

$$\left\langle \boldsymbol{m}^2(\psi) \right\rangle = \frac{4\pi}{3V}\left\langle \sum_{ij} \mu_i \mu_j \theta(\psi - \psi_{ij}) \right\rangle \tag{23}$$

where θ is the Heaviside function. A convenient choice is $\psi = \pi/2$ in which case

$$\left\langle \boldsymbol{m}^2(\psi) \right\rangle = \frac{\varepsilon - 1}{3\varepsilon} + \frac{2}{3\pi}\frac{(\varepsilon - 1)^2}{3\varepsilon}. \tag{24}$$

where for large values of ε, the r.h.s. of Eq. 24 is linear in ε.

The hypersphere method, as discussed below, is a very efficient way to sample the configurations of charged or dipolar particles, in particular, because the pair potentials have a closed analytical expression. It was demonstrated to give results identical to those obtained by using the Ewald procedure. The method is certainly an eloquent illustration of the equivalence

between simulations of Coulombic and dipolar systems of finite size when the interactions are properly derived in agreement with the Poisson equation solved for the considered geometry of the simulation cell.

3D Systems with 2D Periodicity

The previous expressions of the energy allow us to perform simulations of periodic dielectric systems on a physically sound basis. They need to be modified for 3D systems of finite extension in one direction and periodic boundary conditions in the two other directions (slab geometry). This question has been studied for both charged and dipolar systems [24–36]. Similarly to the case of 3D fully periodic systems, summation of the series resulting from the Coulomb interactions between the charges and their replicas is the most frequently used method for obtaining the expression of the Coulomb energy, although solving the Poisson equation is another possibility [37].

Several expressions are possible for the energy of dipolar systems with a finite extent ($L_z = h$) in the z spatial direction and periodic in the x and y directions. When the system is enclosed in an oblique simulation box of extension h in the z direction with basis vectors $\boldsymbol{t}_1$, $\boldsymbol{t}_2$ in the xy plane the expression obtained in the form of an Ewald sum [34, 38] is given by (with the notation of [38])

$$
\begin{aligned}
E = & \frac{1}{2} \sum_{i,j=1}^{N} {\sum_{\boldsymbol{n}}}' \frac{1}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij}|^3} \Big[\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j \, B(|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij}|) \\
& - 3 \frac{\boldsymbol{\mu}_i \cdot (\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij}) \;\; \boldsymbol{\mu}_j \cdot (\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij})}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij}|^2} \, C(|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij}|) \Big] \\
& + \frac{\pi}{2A} \sum_{i,j=1}^{N} \sum_{\boldsymbol{G}\neq 0} \Big\{ \frac{1}{G} \big[D(G, z_{ij}) (\boldsymbol{\mu}_i^{\parallel} \cdot \boldsymbol{G}) (\boldsymbol{\mu}_j^{\parallel} \cdot \boldsymbol{G}) \\
& - \mathrm{i} (\mu_i^z (\boldsymbol{\mu}_j^{\parallel} \cdot \boldsymbol{G}) + \mu_j^z (\boldsymbol{\mu}_i^{\parallel} \cdot \boldsymbol{G})) E(G, z_{ij}) - \mu_i^z \mu_j^z F(G, z_{ij}) \big] \exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{\varrho}_{ij}) \Big\} \\
& + \frac{2\alpha\sqrt{\pi}}{A} \sum_{i,j=1}^{N} \mu_i^z \mu_j^z \exp(-\alpha^2 z_{ij}^2) - \frac{2\alpha^3}{3\sqrt{\pi}} \sum_{i=1}^{N} |\boldsymbol{\mu}_i|^2 ,
\end{aligned}
\tag{25}
$$

where A is the surface area and $\boldsymbol{t}$ a 2×2 matrix with columns corresponding to the vectors $\boldsymbol{t}_1$, $\boldsymbol{t}_2$; $\boldsymbol{n}$, $\boldsymbol{\varrho}_{ij}$ and $\boldsymbol{\mu}_i^{\parallel}$ are bidimensional vectors with components n_x and n_y, $x_j - x_i$ and $y_j - y_i$, and μ_i^x and μ_i^y, respectively. The functions $B(r)$ and $C(r)$ are identical to those given in Eqs. 6 and 7; $|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij}| = \{(x_j - x_i +$

$(\boldsymbol{t}\cdot\boldsymbol{n})_x)^2 + (y_j - y_i + (\boldsymbol{t}\cdot\boldsymbol{n})_y)^2 + (z_j - z_i)^2\}^{1/2}$ and

$$D(G,z) = \exp(Gz)\,\mathrm{erfc}(G/(2\alpha) + \alpha z) + \exp(-Gz)\,\mathrm{erfc}(G/(2\alpha) - \alpha z) \tag{26}$$

$$\begin{aligned} E(G,z) &= \frac{\partial D(G,z)}{\partial z} \\ &= G(\exp(Gz)\,\mathrm{erfc}(G/(2\alpha) + \alpha z) - \exp(-Gz)\,\mathrm{erfc}(G/(2\alpha) - \alpha z)) \\ F(G,z) &= \frac{\partial^2 D(G,z)}{\partial z^2} \\ &= -\frac{4\alpha G}{\sqrt{\pi}}\exp(-G^2/(4\alpha^2) - \alpha^2 z^2) + G^2 D(G,z)\,. \end{aligned}$$

An equivalent expression based on the Lekner summation procedure has been derived [16, 39, 40], but is too lengthy to be reproduced here.

3D Systems with 1D Periodicity

For a system with box sizes h_x, h_y and L_z, periodic only in the z direction, the Ewald summation method gives the following expression for the energy [41, 42]

$$\begin{aligned} E = &\frac{1}{2}\sum_{i,j=1}^{N}{\sum_{\boldsymbol{n}}}' \frac{1}{|\boldsymbol{\varrho}_{ij}, \boldsymbol{z}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|^3}\Bigg[\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j B(|\boldsymbol{\varrho}_{ij}, z_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|) \\ &- 3\,\frac{[\boldsymbol{\mu}_i\cdot(\boldsymbol{\varrho}_{ij}, z_{ij} + \boldsymbol{t}\cdot\boldsymbol{n})][\boldsymbol{\mu}_j\cdot(\boldsymbol{\varrho}_{ij}, z_{ij} + \boldsymbol{t}\cdot\boldsymbol{n})]}{|\boldsymbol{\varrho}_{ij}, z_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|^2} C(|\boldsymbol{\varrho}_{ij}, z_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|)\Bigg] \\ &+ \frac{2\pi}{V}\sum_{i,j=1}^{N}\sum_{G_z\neq 0, G_x, G_y} \mathrm{e}^{-G^2/(4\alpha^2)}\frac{(\boldsymbol{\mu}_i\cdot\boldsymbol{G})\,(\boldsymbol{\mu}_j\cdot\boldsymbol{G})}{G^2}\,\exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}_{ij}) \\ &+ \frac{1}{L_z}\sum_{i,j=1}^{N}{}^{*}\,\boldsymbol{\mu}_i^{\parallel}\cdot\boldsymbol{\mu}_j^{\parallel} P(|\boldsymbol{\varrho}_{ij}|) - 2\,(\boldsymbol{\mu}_i^{\parallel}\cdot\boldsymbol{\varrho}_{ij})\,(\boldsymbol{\mu}_j^{\parallel}\cdot\boldsymbol{\varrho}_{ij}) Q(|\boldsymbol{\varrho}_{ij}|) \\ &+ \frac{\alpha^2}{L_z}\sum_{i,j=1}^{N}\boldsymbol{\mu}_i^{\parallel}\cdot\boldsymbol{\mu}_j^{\parallel}\delta(\boldsymbol{\varrho}_{ij}) - \frac{2\alpha^3}{3\sqrt{\pi}}\sum_{i=1}^{N}|\boldsymbol{\mu}_i|^2\,, \end{aligned} \tag{27}$$

where

$$\begin{aligned} P(|\boldsymbol{\varrho}_{ij}|) &= \frac{1 - \exp(-\alpha^2\boldsymbol{\varrho}_{ij}^2)}{\boldsymbol{\varrho}_{ij}^2} \\ Q(|\boldsymbol{\varrho}_{ij}|) &= \frac{1 - \exp(-\alpha^2\boldsymbol{\varrho}_{ij}^2)}{\boldsymbol{\varrho}_{ij}^4} - \alpha^2\frac{\exp(-\alpha^2\boldsymbol{\varrho}_{ij}^2)}{\boldsymbol{\varrho}_{ij}^2} \end{aligned} \tag{28}$$

$\boldsymbol{n} = (0, 0, n_z)$ and the star * indicates that $\boldsymbol{\varrho}_{ij} \neq 0$.

In this formula, the sums on G_x and G_y correspond to continuous Fourier transforms on the reciprocal variables k_x and k_y of x and y, respectively, writ-

ten as Riemann discrete integrals following the scheme proposed in [43]. The integration steps dk_x and dk_y are chosen equal to $2\pi a_x/h_x$ and $2\pi a_y/h_y$ where the numbers a_x and a_y have values between 0 and 1 which determine the fineness of the 2D integration grid. Defining $L_x = h_x/a_x$ and $L_y = h_y/a_y$, one has $\mathrm{d}k_x = 2\pi/L_x$ and $\mathrm{d}k_y = 2\pi/L_y$ and for any function $f(k_x, k_y, G_z)$ with $k_x \equiv G_x = 2\pi n_x/L_x$ and $k_y \equiv G_y = 2\pi n_y/L_y$ one can write

$$\frac{1}{L_z}\sum_{G_z\neq 0}\int \mathrm{d}k_x \mathrm{d}k_y \mathrm{e}^{\mathrm{i}(k_x x + k_y y + G_z z)} f(k_x, k_y, G_z) \simeq \frac{4\pi^2}{L_x L_y L_z}\sum_{G_z\neq 0, G_x, G_y} \mathrm{e}^{\mathrm{i}(G_x x + G_y y + G_z z)} f(G_x, G_y, G_z) \tag{29}$$

In Eq. 28, V is equal to $L_x L_y L_z$.

The Lekner expression is derived in [33, 44].

2D Systems

Monolayers of dipolar particles are relevant for many experimental systems of molecules with 3D dipoles located on planar surfaces and at interfaces. The Ewald expression of the dipolar energy [45] of such a system is easily derived by taking the limit $z = 0$ in Eqs. 25 and 26

$$\begin{aligned} E = &\frac{1}{2}\sum_{i,j=1}^{N}{\sum_{\boldsymbol{n}}}' \frac{1}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|^3}\Bigg[\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j\, B(\alpha|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|) \\ &- 3\,\frac{\boldsymbol{\mu}_i\cdot(\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n})\;\; \boldsymbol{\mu}_j\cdot(\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n})}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|^2}\, C(\alpha|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|)\Bigg] \\ &+ \frac{\pi}{A}\sum_{i,j=1}^{N}\sum_{\boldsymbol{G}\neq 0}\frac{\mathrm{erfc}(G/2\alpha)}{G}(\boldsymbol{\mu}_i\cdot\boldsymbol{G})\,(\boldsymbol{\mu}_j\cdot\boldsymbol{G})\exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{\varrho}_{ij}) \\ &+ \frac{\pi}{A}\sum_{i,j=1}^{N}\sum_{\boldsymbol{G}\neq 0}\mu_i{}^z\mu_j{}^z\left[\frac{2\alpha}{\sqrt{\pi}}\exp(-G^2/4\alpha^2) - G\mathrm{erfc}(G/2\alpha)\right]\exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{\varrho}_{ij}) \\ &+ \frac{2\alpha\sqrt{\pi}}{A}\sum_{i,j=1}^{N}\mu_i^z\mu_j^z - \frac{2\alpha^3}{3\sqrt{\pi}}\sum_{i=1}^{N}|\boldsymbol{\mu}_i|^2\,. \end{aligned} \tag{30}$$

The formula based on the Lekner summation procedure has been obtained by Grzybowski and Bródka [39].

2D System with 1D Periodicity

The Ewald sum for 2D systems periodic only in one direction (x direction) has been derived in [46]

$$
\begin{aligned}
E = & \frac{1}{2} \sum_{i,j=1}^{N} {\sum_{\boldsymbol{n}}}' \frac{1}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|^3} \Bigg[\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j \, B(\alpha|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|) \\
& - 3 \frac{\boldsymbol{\mu}_i \cdot (\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}) \;\; \boldsymbol{\mu}_j \cdot (\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n})}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|^2} \, C(\alpha|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}|) \Bigg] \\
& + \frac{1}{L_x} \sum_{i,j=1}^{N} \sum_{\boldsymbol{G}} \Big\{ J_0(y_{ij}, G) \, (\mu_i^y \mu_j^y + \mu_i^z \mu_j^z) \\
& - 2 J_1(y_{ij}, G) \, \mu_i^y \mu_j^y y_{ij}^2 \Big\} \exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}_{ij}) \\
& + \frac{1}{2L_x} \sum_{i,j=1}^{N} \sum_{\boldsymbol{G}\neq 0} \Big\{ J_{-1}(y_{ij}, G) \, G^2 \, \mu_i^x \mu_j^x \\
& + \mathrm{i}\, 2 J_0(y_{ij}, G) \, G \, (\mu_i^x \mu_j^y + \mu_i^y \mu_j^x) y_{ij} \Big\} \exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}_{ij}) - \frac{2\alpha^3}{3\sqrt{\pi}} \sum_{i=1}^{N} |\boldsymbol{\mu}_i|^2
\end{aligned}
\tag{31}
$$

where $\boldsymbol{\varrho}_{ij}$ is a 2D vector and $\boldsymbol{\mu}_j$ a 3D vector, B and C are defined in Eqs. 6 and 7. Here $\boldsymbol{n} = (n_x, 0)$ and $\boldsymbol{G} = (2\pi n_x / L_x, 0)$ with L_x the system dimension in the x direction. The functions $J_\nu(y, G)$ are given by

$$
J_\nu(y, G) = \int_0^{\alpha^2} s^\nu \mathrm{e}^{-sy^2 - G^2/(4s)} \, \mathrm{d}s \, . \tag{32}
$$

A convenient way to evaluate $J_\nu(y, G)$ is described in [47]. One can note that if, in the functions J_ν, $\exp(-sy^2)$ is written as a Fourier transform $\int \mathrm{d}k_y \exp(\mathrm{i}k_y y - k_y^2/(4s))(1/\sqrt{4\pi s})$, it is possible to express the contributions of the reciprocal k-space terms in Eq. 31 in a way similar to Eq. 27 by using a modified form of Eq. 29.

Slab in a Dielectric Medium

When a system of finite extension ($L_z = h$) in one spatial direction is confined by dielectric media of dielectric constant ε, the Ewald and Lekner expressions (cf. Eq. 25) for charged or dipolar systems are no longer applicable. For charged systems, each charge and its periodic replica generate an infinite set of image charges. The series which represents the interaction between

charges, replica charges and image charges is a double series,

$$E_C = \frac{1}{2}\sum_{i,j=1}^{N}{\sum_{\boldsymbol{n}}}'\sum_{m} \frac{q_j q_i \Delta^{|2m|}}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij} + 2mh|} + \frac{q_j q_i \Delta^{|2m+1|}}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_i + z_j + (2m+1)h|} \tag{33}$$

where the prime means that for $\boldsymbol{n} = 0$ and $m = 0$ the term $i = j$ is excluded in the sum, $\Delta = -(\varepsilon - 1)/(\varepsilon + 1)$ and

$$|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij} + 2mh| = \{(x_j - x_i + (\boldsymbol{t}\cdot\boldsymbol{n})_x)^2 + (y_j - y_i + (\boldsymbol{t}\cdot\boldsymbol{n})_y)^2 + (z_j - z_i + 2mh)^2\}^{1/2} . \tag{34}$$

The z origin is at mid-distance from the two interfaces. Similarly to the case of systems periodic in the three spatial directions, the series can be computed by multiplying the series terms by a convergence factor; the resulting expression for the energy is [28, 29]

$$\begin{aligned} E_C = {} & \frac{1}{2}\sum_{i,j=1}^{N}{\sum_{\boldsymbol{n}}}' q_j q_i \Bigg\{ \sum_m \Delta^{|2m|} \frac{\operatorname{erfc}(\alpha|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij} + 2mh|)}{|\boldsymbol{\varrho}_{ij} + \boldsymbol{t}\cdot\boldsymbol{n}, z_{ij} + 2mh|} \\ & + \sum_m \Delta^{|2m+1|} \frac{\operatorname{erfc}(\alpha|\boldsymbol{\varrho}_{ij} + t\cdot n, z_i + z_j + (2m+1)h|)}{|\boldsymbol{\varrho}_{ij} + t\cdot n, z_i + z_j + (2m+1)h|} \Bigg\} \\ & + \frac{\pi}{2A}\sum_{i,j=1}^{N} q_j q_i \sum_{\boldsymbol{G}\neq 0} \frac{\exp(\mathrm{i}G\cdot\boldsymbol{\varrho}_{ij})}{G} \Bigg\{ \sum_m \Delta^{|2m|} D(G, z_{ij} + 2mh) \\ & + \Delta^{|2m+1|} D(G, z_i + z_j + (2m+1)h) \Bigg\} \\ & - \frac{\sqrt{\pi}}{A}\sum_{i,j=1}^{N} q_j q_i \sum_m \Bigg\{ \frac{\Delta^{|2m|}}{\alpha} \exp(-\alpha^2 (z_{ij} + 2mh)^2) \\ & + \frac{\Delta^{|2m+1|}}{\alpha} \exp(-\alpha^2 (z_i + z_j + (2m+1)h)^2) \Bigg\} \\ & - \frac{2\pi}{A}\sum_{i,j=1}^{N} q_j q_i \Bigg\{ z_{ij} \operatorname{erf}(\alpha z_{ij}) \\ & - \sum_{m=1,\infty} \Delta^{2m} \big[(z_{ij} + 2mh)\operatorname{erfc}(\alpha(z_{ij} + 2mh)) \\ & + (-z_{ij} + 2mh)\operatorname{erfc}(\alpha(-z_{ij} + 2mh)) \big] \end{aligned} \tag{35}$$

$$- \sum_{m=0,\infty} \Delta^{2m+1} \Big[(z_i + z_j + (2m+1)h)\mathrm{erfc}(\alpha(z_i + z_j + (2m+1)h)) + (-z_i - z_j + (2m+1)h)\mathrm{erfc}(\alpha(-z_i - z_j + (2m+1)h)) \Big] \Big\}$$

$$- \frac{\alpha}{\sqrt{\pi}} \sum_{i=1}^{N} q_i^2$$

where $D(G, z)$ is defined in Eq. 26.

Other expressions have been proposed by Smith [29], who also discusses their efficient computation. An expression for the energy of dipolar systems can certainly be derived, but is not available in the literature.

The potentials due to a charge or a dipole located between two dielectric media with planar interfaces and different dielectric constants are given in [48]. Although the derivation concerns only the case of a system infinite in the x and y directions it nevertheless can provide a first step in the derivation of energy expressions of charged or dipolar periodic systems.

Polarization effects, which induce fluctuating dipoles in addition to permanent dipoles on polarizable atoms or molecules, do not produce essential modifications to the energy expressions of pure permanent dipolar systems when written in the context of the Ewald method. Generalization of the Ewald expressions for mixtures of molecules possessing charges, dipoles, quadrupoles and polarizabilities can be found in [49, 50] and [51].

In most Q2D experimental realizations of dipolar systems, the dipoles interact by the 3D dipole-dipole potential. However, systems of dipoles interacting by the potential, derived from the solution of the 2D Poisson equation $q \log r$, ($p = 2$ in Eq. 1), are of interest too. The expressions of the pair interaction which must be used if the system has periodic boundary conditions have been determined both by the methods of series summation and the Poisson equation [52, 53].

Linear chains of fixed dipolar particles have been extensively studied in the literature, but this topic is out of the scope of this review.

Numerical Implementation

The efficient implementation and programming of the above described energy expressions are the main problem in the simulation of charge or dipole systems. When electric dipoles or possibly electric multipoles are described by an intra-molecular charge distribution the relevant energy expressions are those of the charge systems where the Coulombic intra-molecular contributions are omitted.

For systems of point dipoles or higher multipoles the difficulties which must be overcome to realize efficient codes are similar to those encountered

for the charge systems; the proposed solutions are similar for the two types of systems.

Clearly when the Ewald energy expressions are used, an adequate choice of the parameter α must be made to minimize the number of terms needed to compute the series in r- and k-spaces. In most implementations, α is chosen so that in the computation of the series in r-space, only the term $n = 0$ needs to be retained. The number of terms to be considered in the series in k-space is determined so that the numerical error on the estimate of the dipolar energy is equal or lower than a given value. When α is chosen according to these conditions, it was shown that the computation time of the Ewald sum scales as $N^{3/2}$ [54]. Studies leading to a precise determination of α values associated with a controlled numerical error on the energy or electric field values have been made for the charge systems [54, 55] and recently completed in detailed work [56] for the dipolar particle system. The number of terms needed to estimate accurately the Coulomb energy by using Lekner sums has also been discussed [14, 57]. For most of the relative positions of a particle pair, the value of the pair interaction is very accurately estimated with only a few terms in the Lekner sum. However, when the vector distance between a pair of particles is approximately parallel to the xy plane, the functions K_ν are divergent and a large number of terms must be considered to reach a precise estimate of the pair energy. Numerical methods allowing to avoid such a shortcoming have been described in [15, 39, 57].

It is worth remarking that in a simulation using the S_3 geometry the Coulombic or dipolar energy is computed with an accuracy depending only on the precision of the floating arithmetic of the computer, since the dipole-dipole pair potential has a closed analytical form. In addition, in the S_3 geometry, the computation of an interparticle distance needs less arithmetic operations than in the periodic system where such a distance is calculated by using the minimum image convention. However, in spite of these advantages, the time required for the energy computation in the S_3 geometry increases as N^2, and the simulations in S_3 are competitive with those based on the use of Ewald sums, where this time scales as $N^{3/2}$, and only for systems of moderate sizes ($N < 5000$ particles).

Large Systems

As an adequate choice of the α parameter in the Ewald sums allows the computation time of the contribution to the energy coming from the series in r-space to scale linearly with N, several methods have been devised to reduce the computation time of the series of terms in k-space. The latter is a Fourier transform depending on the particle positions, which, obviously in continuous systems, have disordered locations. The disorder of the particle positions seems to preclude the use of a fast Fourier transform (FFT) algorithm for the computation of the k-space series. Adequate numerical schemes have, however, been devised to overcome this difficulty. In the particle-particle

particle-mesh (P3M) scheme [58–63], the first step is to remark that the series in k-space is equivalent to the Coulombic energy of Gaussian charge distributions centred at the particle positions. Clearly, this energy can be determined from the solution of the Poisson equation associated with the charge density (sum of these Gaussian distributions). When this density is assigned to a regular grid, its Fourier transform, computed by a FFT, gives the solution of the Poisson equation. Elaborate assignments of the charge density have been proposed to minimize the computation time and numerical error associated with the finite size of the 3D grids used in practice in the simulations [64, 65]. The use of the FFT algorithm to compute the k-space series is further made possible by approximating the complex exponentials $\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r}_i)$ by linear combinations of complex exponentials calculated at grid points close to the particle positions $\boldsymbol{r}_i$ [66]. Detailed descriptions of such schemes known as particle-mesh Ewald (PME) [59] and smooth particle-mesh Ewald SPME [62] have been made for systems of charged, dipolar and polarizable particles [65]. For N_G grid points, the computation time of the k-space series is of the order $N_G \log(N_G)$; if the number of grid points is assumed proportional to N (an acceptable estimate for dense fluid or solid systems), the computation time of the Ewald sum varies with the system size as $N \log(N)$. Detailed implementations of the P3M, PME and SPME methods, needed to achieve the computation of energy and forces with a given accuracy, are described in [67, 68].

The numerical methods, known as the Fast Multipole Method (FMM) [69–72] and the Cell Multipole Method (CMM) [73–76], are particularly suitable for systems of large numbers of charges and/or dipoles. In brief, they are based on a division of the volume of the simulation cell into sub-volumes. The contribution to the energy of the interaction between charges within the same sub-volume is computed exactly by direct summation and that due to interactions between charges in different sub-volumes by considering that the potential acting on a charge in a sub-volume m is the sum of the potentials associated with the multipole expansion of the charge distribution inside the sub-volumes $n \neq m$. The decrease of the computation time results from the important reduction of the number of distances between particles needed to compute the Coulombic or dipolar energy. Thus, the contribution to the energy of a particle in a sub-volume m involves the distances between the particles inside this sub-volume and those between this particle and the centres of the other sub-volumes, the points around which the multipole expansions are made. The cost of the energy computation increases as N when N is sufficiently large. The main difference between the FMM and CMM methods lies in the way the multipole expansions are implemented using either spherical harmonics or Taylor expansions. The question of compatibility of the multipole method with periodic boundary conditions has been considered in several publications [72, 77–79]. A detailed description of the practical use of the multipole method for systems of dipolar and polarizable particles

is given in [80, 81]. Using the same idea of a division of the potential acting on a charge in two parts due to the neighbouring and distant charges, respectively, Sperb and Strebel have proposed the MMM method which is based on an energy expression derived in [14]. A remarkable use of the FMM method for large 2D dipolar Ising systems has been made by Stoycheva and Singer [82] in their study of stripe melting in magnetic monolayers. We finally mention that the previous FMM methods have been considered for 2D periodic charged and dipolar systems based on the $\log r$ potential [52] by Christiansen et al. [74].

Partially Periodic Systems

The case of 3D systems finite in one or two spatial directions and periodic in the other directions (slab geometry) needs a special treatment. As seen from Eqs. 25 and 27 the number of arithmetic operations required for the computation of the k-series in the Ewald expression of the energy are of the order N^2 multiplied by the number of terms retained for the numerical estimate of the series. The cost of such a computation is generally prohibitive even for systems of small sizes. Algorithms have been developed to overcome this problem and maintain the computational cost of the k-series proportional to N. The method proposed by Hautman and Klein [83] for systems of finite size in the z direction is based on an expansion of $1/r$ in powers of z/s, where z and s are the components of $\boldsymbol{r}$ perpendicular and parallel to the slab surfaces, and the contribution to the energy of each term of the expansion is obtained by computing the sum on the charge replicas by an Ewald method. For narrow slabs, an accurate value of the energy can be obtained by retaining only a few terms in the expansion of the r- and k-space contributions; it is possible to rewrite the latter in a form corresponding to a computational cost proportional to N.

The Lekner expressions of the energy of systems of charges or dipoles periodic in one or two directions [39] contain only terms in r-space, and therefore seem particularly adequate for an efficient calculation. However, it is also supposed that use of the Lekner sum, as was mentioned for the 3D systems, overcomes the numerical instabilities arising from the divergence of the functions K_0 and K_1 for small values of their arguments [57]. An extension of the MMM method to the systems with a slab geometry is discussed and checked in [84].

A simple recipe to perform calculations for a system of height h in the z direction and periodic in the x and y directions, with sides L_x and L_y, is to include it in the central part of a volume of sides L_x, L_y and L_z with periodic boundary conditions in the three spatial directions. If L_z is chosen much larger than h, it is expected that the effect of the replicas in the z direction will not affect the calculated physical properties of the system. Such a conclusion has been obtained in [85]. However, if a large value of L_z has to be used, the number of terms necessary for a precise estimate of the k-space series also increases. A method to

reduce this inconvenience, is to add a surface term proportional to the square of the z component of the total dipole moment of the system to the shape independent part of the 3D Ewald energy as proposed in [86, 87]. This term is obtained by performing the sums in Eq. 5 in a slab-wise fashion. This approach, has been recently improved in [88, 89] by an exact evaluation of the interaction between replicated layers which has to be subtracted. In this way the energy of systems with a slab geometry can be computed accurately by using the 3D Ewald sum with an acceptable value of L_z.

An efficient numerical method, developed and used in [43, 90–92] for the computation of the Coulomb energy of 3D systems with 2D periodicity has been shown in [93] practically equivalent to the above described methods based on the use of the Ewald sum for 3D systems with 3D periodicity.

Polarizable Systems

In numerical simulations of systems of N polarizable dipolar particles the induced dipoles must be computed by solving a set of N linear equations depending on the positions and orientations of the N dipoles of the system. The solution of this set of equations requires of the order of N^2 operations and deteriorates considerably the efficiency of MC simulations since, in principle, for each MC elementary move involving, for instance, the displacement of one particle, the N linear equations have to be solved. The validity of numerical procedures allowing us to overcome this problem is discussed in [94]. The procedures are based on the choice of an adequate cut-off of the interparticle distances such that the computation of the induced dipole of a displaced particle depends only on the positions and orientations of dipoles located at a distance of the trial particle position smaller than the cut-off.

3 Simple Dipolar Fluids

The systems of particles with permanent or induced electric or magnetic dipoles have physical properties induced by the specific characters of the dipole-dipole interaction. One of these properties is the dielectric constant which, in simulations, is obtained from the average value of the fluctuations of the dipole moment $\langle \boldsymbol{M}^2 \rangle$ by using Kirkwood or Onsager relations according to the boundary conditions [95]. An accurate estimate of the dielectric constant requires some caution [96, 97]. For fluids of particles with a permanent dipole moment comparable to that of water, the computation of the dielectric constant is best performed with the use of the Ewald potential with tin foil boundary conditions and requires at least 500 particles and a number of time steps larger than 10^6 if MD runs are performed. Systems where the surface term E_P (Eq. 11) is associated to the Ewald energy, do not present a noticeable advantage over those with tin foil boundary conditions since for the latter the

uncertainty on the dielectric constant value is minimal for a given statistical error on the calculated average of $\langle \boldsymbol{M}^2 \rangle$. Determination of the dielectric constant from computing the linear response of the total dipole moment to a weak external field is also possible, but this approach is not significantly more efficient than the previous method [97].

Systems of dipolar molecules are common, and their phase diagrams, generally, do not differ qualitatively from those of simple fluid models where the particles interact by a spherically symmetric van der Waals potential. For instance, it is expected that a small dipole moment influences the location of the liquid-gas (l-g) critical point but does not preclude this first order transition. It was realized, mainly on the basis of numerical simulations, that the phase diagram of systems where the dipole-dipole interaction is the dominant interaction differs largely from the standard scheme of successive gas-liquid and fluid-solid transitions. In addition strong dipolar interactions seem to be able to generate fluid or solid ferroelectric phases. The properties of simple systems of dipolar particles are presented and discussed below.

3.1 Phase Diagram

The simplest continuum dipolar system is certainly point dipoles embedded at the centre of a spherical hard (DHS) or soft ($1/r^{12}$ or truncated and shifted LJ potential) (DSS) core to prevent divergence of the dipolar interaction at small separations. Yet the dipolar interaction presents two peculiarities which make the system not so simple after all: a slow spatial decay ($\sim 1/r^3$) and a highly directional angular dependence. A consequence of the long range behaviour is that the dipole interaction may create shape dependent internal demagnetization fields whose strength depends on the boundary conditions [98]. The demagnetization field is at the origin of domain formation in high density dipolar systems. A proof of the independence of the free energy on the shape and boundary conditions in zero applied field and of the thermodynamic limit has been given recently by Banerjee, Griffiths and Widom [98]. The second characteristic feature of the dipolar interaction, its strongly directional angular dependence, explains the tendency for chain formation in low density DHS and DSS models. Two close dipoles aligning "head-to-tail" have indeed a much more favourable energy ($-2\mu^2/\sigma^3$) than arranging side by side with antiparallel ($-\mu^2/\sigma^3$) orientations.

An intriguing question is whether DHS and DSS models can show l-g coexistence or whether chain formation precludes formation of a liquid. A simple argument seems in favour of the occurrence of a l-g transition. It is based on the estimate of the effective interaction, at large distance, between dipolar particles which behaves like an attractive van der Waals potential ($-1/r^6$). But this estimate which results from the average of the dipole-dipole interaction weighted by its Boltzmann factor, is unconvincing since it neglects the effect

of strong short ranged correlations between dipole moments. On the basis of systematic MC studies of the DHS model, it became quite clear that the l-g transition was not present in the expected domain of densities and temperatures. MC simulations performed in the NpT and Gibbs ensembles for temperatures between $T^* = 0.222$ and 0.18 and densities $0.1 \leq \rho^* \leq 0.4$ [22] and in the NVT ensemble, mainly along the isotherm $T^* = 0.0816$ and densities $0.02 \leq \rho^* \leq 0.3$ [99, 100], did not find any evidence for a l-g transition.

As apparent from 3D simulations of DHS [99, 100] and DSS [101–105] small chains (dimers, trimers) start forming when the dipolar energy exceeds the thermal energy ($\lambda \sim 3$). By increasing the dipole moment (or lowering the temperature) the chains grow and an equilibrium state establishes consisting of a distribution of chains, rings or more complex structures of different lengths, breaking and reforming, with relative concentrations depending on density and temperature.

Chain formation is a slow process, especially at low density. Two chains obviously like to attach if two end dipoles become closely aligned head-to-tail, rings can form from two chains of favourable relative positions and dipole orientations or if the two ends of the same (large) chain meet with favourable orientation of their dipole moments. In longer chains an end particle may attach to an interior particle thereby forming a ring and a chain that subsequently detaches; two chains coming into contact can exchange branches at "Y" or "X" junctions; two rings can fuse into a larger ring or a larger ring can form from two rings [106] etc.... As chains assemble, at large dipole moments, spontaneous breaking will occur only if neighbouring chains come close and approach one another with favourable relative orientations [107]. In view of these mechanisms through which evolution of the system takes place it is clear that in MC simulations cluster moves [108, 109], translating and rotating cluster entities as a whole, in addition to single particle moves, will greatly enhance sampling of phase space at low density. The modification of the acceptance ratio in MC simulations with cluster moves is found in [110, 111] for a wide class of cluster rules. In particular, detailed balance requires that a trial cluster move be rejected if the connectivity of the cluster changes in this move.

At low densities ($\rho^* \leq 0.01$) [99, 112] the system of 3D DHS consists of a mixture of short chains and rings. As the density increases at low temperature, (cf. results along the isotherm $T^* = 0.1322$ [112]), the fraction of particles in chains and the average cluster sizes increase; at $\rho^* \simeq 0.06$ there are no more ring like clusters but a network of long chains which span the simulation cell. At higher density ($\rho^* \geq 0.2$) the network structure disappears and the system is more like a normal liquid. Eventually the system gets ferroelectric [112]. These structural changes are also apparent in the low wave vector behaviour of the structure factor [113]. A rough sketch of the different phases evidenced in DHS is given in Fig. 1.

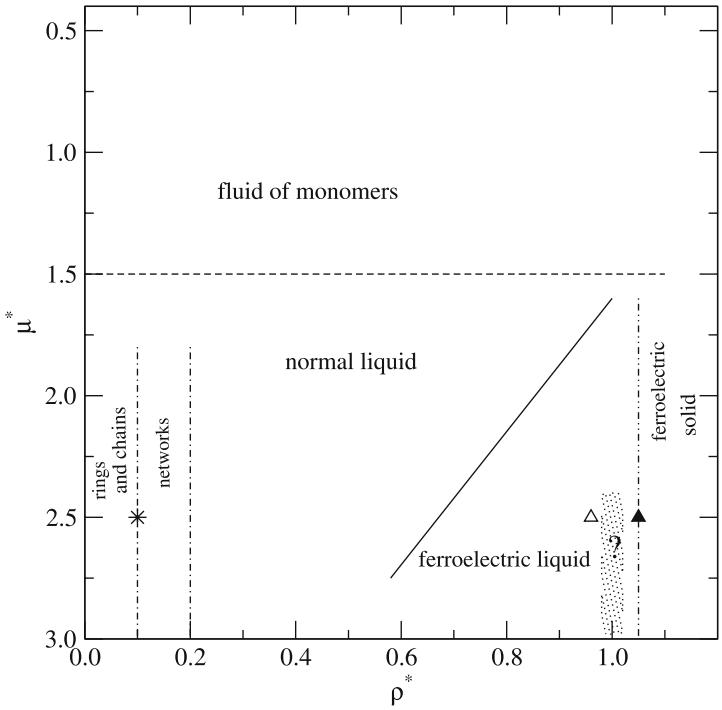

Fig. 1 Sketch of the phase diagram of 3D DHS in the temperature-density plane. The reduced temperature is related to the reduced dipole moment by $T^* = 1/\mu^{*2}$. At high temperature ($\mu^* \leq 1.5$) the system consists mainly of monomers. At lower temperatures and low density the DHS particles are associated in chains and rings. Upon increase of the density the average cluster sizes increase, the rings unbind in favour of chain-like segments and the structure changes to that of a network spanning the (finite) simulation box. At densities $\rho^* \geq 0.2$ each particle is coordinated with a whole shell of particles and the network structure breaks down: the system is more like a normal liquid [99, 113]. Further compression gives rise to a ferroelectrically ordered state [99, 102, 135]. Between the ferroelectric solid and the ferroelectric liquid a small region (*dotted region*) may exist with ferroelectric columnar ordering of the particles [102, 134]. The star marks the critical point of a condensation transition conjectured in [112, 113]. The triangles indicate the limits of mechanical stability of the bct (*open triangle*) and fcc (*filled triangle*) solid phases [135]

A detailed study of the isolated chains or chain network requires a working definition of a chain which can be based on either a steric or energetic criterion (e.g. two spheres or particles are bound if their distance is smaller than a predetermined distance or their potential energy lower than a predetermined value E_c) [99, 100]. As shown by Tavares et al. [114] results will depend somewhat on the precise choice of the criterion. For instance in 3D DHS, the average chain length at $\rho^* = 0.05$ and $\mu^* = 2.75$ is $\sim 4.8\sigma$ for a threshold energy $E_c = -1.4\mu^{*2}$ and $\sim 3.2\sigma$ for $E_c = -1.5\mu^{*2}$. Using a value $E_c = -1.4\mu^{*2}$

a typical average chain length at $\rho^* = 0.02$ is $\sim 8\sigma$ and $\sim 24\sigma$ at $\rho^* = 0.10$ ($\mu^* = 3.5$) while persistence lengths at these densities and dipole couplings are $\sim 4.3\sigma$ and $\sim 7\sigma$, respectively.

The structural characteristics of a Q2D dipole system differ from those in 3D by a larger tendency for ring formation and defect clusters of more complex topology as well as by the absence of a ferroelectric phase at high density. At large density the Q2D dipolar hard spheres rather organize in vortex-like structures with no net polarization [7, 106, 115, 116]. Interestingly similar patterns have been observed in a vertically vibrated monolayer of magnetic spheres [117].

Due to limits on system size and lack of efficient sampling, a complete analysis of the chain structure in 3D dipolar systems has not been realized. Only in 2D, could systems of sufficiently large size (5776 particles) be investigated so far allowing a quantitative description of chain and ring length distributions and conformational properties [7, 106]. A detailed comparison of the length distributions of the simulated dipolar chains and rings with those predicted by equilibrium polymer theory shows that the structure of the Q2D dipolar system is, to a good approximation, the same as that of 2D equilibrium polymers [7, 106].

The compactness of chains or rings, considered as clusters of associated particles, can be measured by the radius of gyration R_g which should scale as $N \sim R_g^d$ where N is the number of monomers in a cluster and d the cluster fractal dimension. The simulation results for strongly interacting Q2D dipolar HS show that small clusters, $N \leq 10$ for chains and $N \leq 40$ for rings, scale with $d \sim 1$ indicating that they behave as rigid objects while larger clusters have the conformation of 2D self-avoiding walks ($d \sim 4/3$) [106]. This contrasts with recent experiments with magnetized steel particles thermalized by vibrations where clusters with a small number of particles appear more compact ($d \sim 2$) while larger clusters become more extended ($d \sim 1.4$) [118]. Network states similar to those seen in the simulations are obtained only if the high temperature gas is rapidly quenched to a low temperature and are conjectured to be metastable states [118].

Linear chain structures resembling those described above could be observed recently in cryo-TEM (transmission electron microscopy) studies of thin films of metallic iron dispersions of sufficiently large size in a zero field [119, 120]. In these experiments the solvent is vitrified by a fast temperature quench allowing the structural arrangement of the arrested particles to be visualized by direct imaging. The formation of linear chains and network like clusters gives strong evidence that in these iron dispersions the magnetic interactions are dominant over the isotropic repulsive and van der Waals interactions stemming from the coating of the magnetic cores by a surfactant layer to prevent irreversible aggregation of the particles. Alignment of the chains along the direction of an externally applied saturating magnetic

field and lateral aggregation of the chains into bigger clusters [119, 120] give further support for a system dominated by dipolar interactions (see below). In the experiments of Butter et al. [119] the onset of chain formation starts roughly at a dipole moment $\mu^* \approx 2.5$ in good agreement with simulation. However, the observed average chain length, 2.49, at the density $\rho\sigma^2 = 0.4$ and $\mu^* \approx 2.4$ is much smaller than the simulation result [106]. Also, surprisingly, no formation of rings, ubiquitous in the simulation studies of Q2D dipolar HS, is found in the samples. In more conventional ferrofluids like oleic-acid-coated magnetite dispersions the dipole interaction is much weaker (lower bulk saturation magnetization) and the structure is believed to be governed by the van der Waals interaction yielding isotropic clusters or drops rather than linear chains [119] in accord with simulations of Q2D Stockmayer fluids [45] or DHS with additional Yukawa attraction [7].

Nature has succeeded in producing even stronger dipolar magnetic particles inside the cells of magnetotactic bacteria [121]. These bacterial magnetite particles have a size (≈ 50 nm) much larger than that of typical laboratory-made magnetite colloids (5–15 nm) and therefore a larger dipole moment ($\lambda \approx 50$). When extracted from the cells and confined to lie on a 2D plane various morphologies including linear chains, rings, rings with attached chains or more compact assemblies are observed [122].

Self-assembly into chains, closed loops, netlike structures or hexagonal superlattices has also been demonstrated in Q2D systems of magnetic microspheres [107] and spherical ε-Co nanocrystals of appropriate size [123] as well as in simulations of elongated molecules with transverse dipole moment (see Sect. 4) [124, 125].

Bounds on the temperature-density range for possible l-g coexistence have been suggested in 3D DHS and DSS systems. Van Leeuwen and Smit [126] have examined, through Gibbs ensemble MC simulations, the change of a coexistence curve of a system with short range potential

$$v_{sr}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \varepsilon_6\left(\frac{\sigma}{r}\right)^{6}\right] \tag{36}$$

by varying the relative strength of the dipolar to attractive (nondipolar) interaction, i.e. interpolating between the Stockmayer fluid model and the DSS model. For $\varepsilon_6 = 1$ the system exhibits l-g coexistence (at least for not too large dipole moments) because the LJ potential does so. Upon decreasing ε_6 to zero the critical temperature and density decrease monotonically and the width of the coexistence curve shrinks. For $\varepsilon_6 < 0.3$ no coexistence could be found anymore. In this domain of density and temperature snapshots of configurations of the system revealed chaining of the particles. Extrapolation of the critical density to $\varepsilon_6 = 0$ yields an upper bound $\rho_c^* \sim 0.03$ for the critical density (provided it exists) of the DSS system. The same behaviour should be observed by increasing the dipolar strength in a Stockmayer potential [127]. As the strength of the dipolar interaction increases the critical

temperature $T_c^* = kT_c/\varepsilon$ increases and the critical density decreases but λ saturates to a value $\lambda_c \sim 25$. For values of μ and T in the vicinity of λ_c chain formation occurs.

An increase of T_c^* is also observed if a polarizability is added to Stockmayer particles, since the main effect of polarizability is to increase, on average, the magnitude of the permanent dipole moment [128]. Similar sensitivity of the l-g transition to the relative strength of the dispersion potential and dipolar interaction occurs in the Q2D Stockmayer fluid [45]. Compared to the 3D Stockmayer fluid the critical temperature is reduced by confinement of the centres of mass in a plane and the coexistence curve is much flatter. In [45] no estimate has been given for the dipole strength at which the l-g transition is expected to be superseded by the chain formation.

A second delimitation of the (T, ρ) region where a l-g transition in DHS could exist comes from simulations of dipolar hard dumbells and spherocylinders [129]. If the hard core of the dipolar particles has a slightly elongated shape in the direction of the dipole moment the head-to-tail configuration is less favoured with respect to the side by side configuration (due to a larger separation of the dipole moments in the direction of the molecular axis) and a l-g transition is restored. Not surprisingly, slightly elongated hard dumbbells and spherocylinders ($L/D \leq 0.28$ at $T^* = 0.12$; L length and D diameter of the molecules) exhibit l-g coexistence even at the smallest elongation ($L/D \sim 0.1$) which could be studied reliably by the Gibbs ensemble and grand canonical MC simulations [129, 130]. At smaller aspect ratios association of the particles into energetically very stable chain structures at low temperature and density reduces dramatically the probability of successful insertion and deletion moves in the Gibbs ensemble and grand canonical MC methods. Extrapolation to the $L/D = 0$ limit would locate the critical temperature of dipolar HS roughly at $T^* \sim 0.18$ with a very low critical density [129].

In view of these results it seemed promising to investigate in more detail the temperature region slightly below $T^* = 0.18$. Camp et al. [112] have performed three independent sets of MC simulations for dipolar hard spheres along the isotherm $T^* = 0.1322$ which give compatible results pointing towards the existence of a "condensation transition between a highly associated vapour phase and a more normal dense liquid phase" [113]. The critical temperature is estimated to be $T^* \approx 0.16$ and the critical density $\rho_c^* \approx 0.1$ [113]. A striking observation is that over the density range covering the different phases the energy and enthalpy remain practically constant in contrast to a normal l-g transition.

Motivated by the simulation results, Tlusty and Safran [131, 132] proposed a mechanism for the phase transition in low density dipolar fluids where competition is brought into play between a low density entropically favourable phase rich in chains (ends) and a higher density energetically favourable phase rich in junctions of three chain ends (Y defects). Not enough simulation data, in particular concerning the formation of Y junctions, is

presently available for 3D DHS to assess the relevance of this mechanism of a phase transition. For Q2D dipolar DHS detailed analysis of the topology of the clusters, of their internal energy and their size distributions suggests that the system undergoes a phase transition from a dilute phase characterized by a number of disconnected clusters to a condensed phase characterized by a network of spanning (macroscopic) clusters that include most of the particles in the system [133].

3.2 Ferroelectric Liquid and Solid

A second surprise, besides that of the absence of a l-g coexistence in low density dipolar hard and soft spheres, is the discovery, by simulation, of an orientationally ordered ferroelectric phase at high density [101–103, 127, 134–136]. This is all the more astonishing as ferroelectric phases have so far been observed only in chiral smectic-C liquid crystals [137] or in undercooled liquid metals [138]. In ferrofluids the high packing fractions at which magnetic ordering occurs in the simulations are presently not attainable. However, the observation of growth of ferromagnetic fluctuations in an iron-nitride sample, at the highest densities that could be investigated ($\rho^* = 0.3$), is suggestive of a ferromagnetic transition [139]. On the other hand, the possibility of the existence of a macroscopically polar liquid has been questioned on the basis of symmetry arguments [140]. It is worth remarking that in experimental systems the values of the particle sizes and dipole moments can present a noticeable polydispersity. Simulations of dipolar HS systems [141] with a Gaussian distribution of the dipole moments, show that the stability of the polarized fluid phase is affected by the distribution width. For identical values of the average dipole moment, the ferroelectric phase can exist for a narrow distribution but will be absent for a large distribution.

Stable ferroelectric nematic phases have been found in high density simulations of DHS [134, 135], DSS [101–103] and Stockmayer particles [127, 136]. So far the transition densities at fixed λ or temperatures (at fixed density) have been estimated only qualitatively by the value at which the polarization order parameter (rounded by finite size effects) $P_1 = \langle(\sum_{i=1}^{N} \hat{\boldsymbol{\mu}}_i \cdot \hat{\boldsymbol{d}})/N\rangle \approx 0.5$ where $\hat{\boldsymbol{\mu}}_i$ is a unit vector and $\hat{\boldsymbol{d}}$ the instantaneous director. A more precise estimate of transition temperatures is presently being attempted [142] through a finite size scaling analysis [143, 144].

Increasing the dipole coupling shifts the transition densities to lower densities. For instance, in dipolar soft spheres (256 particles) magnetic ordering sets in at a density $\rho^* \approx 0.85$ for $\lambda \sim 4$ [103], $\rho^* \approx 0.65$ for $\lambda \sim 6.25$ [102, 103] and $\rho^* \approx 0.55 - 0.60$ at $\lambda \sim 9$ [102, 103]. Transition densities are similar in DHS [113, 135], but are much higher in Stockmayer fluids where a ferroelectric transition exists only for $\lambda \geq 4$ [127, 136] at densities generally larger than the liquid densities at l-g coexistence. MC simulations by Gao and

Zheng [136] predict a narrow low temperature range within which the vapour coexists with a ferroelectric liquid. It is worth stressing that at strong coupling, $\lambda \geq 7$, relaxation of the orientational structure is slow and the system may get trapped for long periods in metastable configurations [102, 103, 142].

By examining the stability of various lattice structures, body-centred-tetragonal (bct), body-centred-orthorhombic (bco), body-centred-cubic (bcc), face-centred-cubic (fcc) and hexagonal close packed (hcp), in isothermal-isobaric MC simulations with variable box lengths, Gao and Zheng [136] conclude that in Stockmayer fluids the bco crystal is the stable phase near the triple point. The latter is obtained from Gibbs free energy calculations in the solid (bco) and liquid phases and is located at a temperature $T^* \approx 1.31$ and a pressure $p^* \approx 0$ ($\mu^* = 2.5$). Groh and Dietrich [145] show that at $T^* = 0$, the Stockmayer model can present hexc (hexagonal phase with slightly contracted c-axis), bco or bct phases, depending on density and dipole moment. For dipole moments $\mu^* \leq 3$ (the values commonly used in simulations) the ground state is bco or hexc, depending on density.

Dipolar soft spheres solidify into a bct crystal at $\rho^* \approx 0.87$ for $\lambda = 6.67$ ($\mu^* = 3$, $T^* = 1.35$) and $\rho^* \approx 0.8$ for $\lambda = 9$ ($\mu^* = 3$, $T^* = 1$) [102]. Inspection of the pair distribution shows that in the ferroelectric liquid the short range spatial correlations are similar to those in the bct crystal [102]. The ground state phase diagram of DSS is quite complex depending on the softness parameter n [145]. For $n = 12$ the sequence of transitions with increasing density is fco (face-centred-orthorhombic)-bct-bco-hexc-bco-fcc for $\mu^* = 1$ [145].

In DHS, at $\mu^* = 2.5$, the bct lattice has been found [135] to be stable over the whole density range considered $\rho^* = 0.95 - 1.20$ but a fcc lattice is also mechanically stable in the region 1.02 – 1.20. Below $\rho^* \approx 1.06$ the fcc lattice transforms into a bco structure. The relative stability of the different phases has, however, not been investigated. In the fcc lattice the polarization is along a [001] direction, and in this direction the dipole moment orientations vary with helical order along lattice planes separated by one-half lattice spacing. In each lattice plane the dipole moments orient nearly parallel making an angle of $\sim 40°$ with the polarization axis [135].

Between the polarized nematic and solid phases a density region seems to exist where dipolar hard and soft spheres form ferroelectric columns with square ordering in the plane perpendicular to the columns [102, 134]. The stability and density range of this phase needs further investigation. Such a columnar phase has also been identified in a system of extended dipoles where the columns order on a hexagonal lattice [146].

Ferroelectric fluid phases have also been observed in simulations of disk-shaped particles with embedded dipole moments along the symmetry axis [147, 148] (see Sect. 4) and in a dipolar hard sphere model carrying two parallel dipole moments displaced equally from the sphere centre [149]. In the latter model antiferroelectric arrangement of the particles is, however, equally preferred at sufficiently large separations (0.3σ) of the two dipole moments.

It is worth stressing that a single ferroelectric domain is obtained in simulations only with the so-called conducting b.c. In this way polarization charges appearing at the system's boundary are absorbed by the conducting medium and there is no depolarization field. This is not the case with vacuum b.c. ($\varepsilon' = 1$) where an unfavourable (positive) depolarization energy is counteracted by the system by splitting into different fully polarized domains with polarization directions such that the total polarization of the system vanishes [102, 135]. Note, however, that the thermodynamic limit must be the same for both boundary conditions [98].

Mixtures of neutral HS and DHS with equal or unequal core diameters undergo demixing when the value of the dipole moment increases (or similarly the temperature decreases). The onset of the demixing depends on the ratio of the core diameters of the two species [141, 150].

3.3 Clusters

The ground state configurations of dipolar clusters have been studied as a function of their size and the strength of the dipolar interaction and compared to those when the interactions between particles are short ranged. A detailed study of the lowest energy configurations adopted by DHS systems with particle number N ranging from 3 to 20 allows us to conclude [151] that linear chains are the most stable arrangement for $N \leq 3$, planar rings for $4 \leq N \leq 14$ and double rings for $14 \leq N \leq 20$. By adding a quadrupole of increasing strength to the DHS, the rings become unstable and compact structures are the typical ground state configurations. If the HS core is replaced by a LJ potential, compact structures are also the most stable configurations when the LJ potential is stronger than the dipole-dipole potential. A small polarizability does not affect the qualitative structure of the ground states, but at larger values, the ring arrangements disappear in favour of chains [151]. The orientational arrangements of lowest energy of 3D dipoles located at fixed positions have been studied in [152]. This study which confirms the results found in [151] determines the minimum energy of finite size crystals where the nearest neighbour distance is equal to the diameter of an associated DHS system. The results show the propensity of dipole orientations for adopting ring or columnar configurations. The orientational ordering in larger clusters of up to 50 Stockmayer particles has been investigated by Lu and Singer [153]. At the dipole moment $\mu^* = \sqrt{3}$, they find that the clusters have an oblate shape with a high degree of circulating orientational order. The effect of an external magnetic field on the stability and breakup of cluster configurations is described in the works of Jund et al. [154] and Kun et al. [155].

The relative influence of the van der Waals and dipolar interactions on the homogeneous l-g nucleation has been determined by ten Wolde et al. [156, 157] for the Stockmayer fluid at $\mu^* = 4$ and temperature $T^* = 3.5$ ($\lambda = 4.6$)

below the critical temperature $T^* = 5.07$. The cluster-size distribution was computed via a biased grand canonical scheme allowing an efficient sampling in the vapour phase of the configurations comprising large clusters (liquid drops) relevant to the nucleation process. The bias is evidently corrected for to estimate the true relative probability of such configurations compared to those where particles are distributed over many small clusters. A main result of this work is that nucleation of the liquid is preceded by the formation of chains. The analysis of particle configurations demonstrates that small clusters containing up to 30 particles are not compact, but correspond to linear or ring arrangements and differ markedly from those observed in systems of particles interacting by LJ potentials [158]. The configurations of larger clusters are compact and of globular type, but their density profiles differ from those occurring in clusters of LJ particles; in particular chains are present in the interfacial region. These simulation data can contribute to explain the discrepancy between the results of classical nucleation theory and the nucleation rates observed in polar liquids.

The influence of an external electric field on the nucleation process has been investigated in the case of the Stockmayer fluid by computing the free energy of formation of clusters in the vapour phase [159]. The free energy is found to increase in the presence of an electric field at fixed supersaturation vapour and to decrease at fixed chemical potential.

3.4 Frozen Systems

As shown above, the ferroelectric phase seems to be the stable phase of dense fluid systems of strongly dipolar particles. When the configurations of the particles are randomly frozen without hindering the dipole reorientations, mean field theory predicts that a stable ferroelectric phase can exist [160]. This possibility has been investigated by numerical simulations of DSS [161, 162] with dipole moments having one, two or three components corresponding to Ising and XY models with long range interaction and the usual DSS system, respectively. For the soft sphere system at $T^* = 10.5$ and $\rho^* = 0.8$, structureless initial configurations were generated except for short range correlations due to the soft cores. After freezing these disordered configurations, the equilibrium arrangements of the three different types of dipoles with $\mu^* = 4$ were computed by constant temperature MD simulations for the XY and DSS models, and by MC simulations for the Ising-like model. In the Ising-like model, spontaneous polarization occurs below $T^* = 25$. For the disordered frozen XY and DSS models no evidence for a ferroelectric state was found upon decreasing the temperature. By contrast, the growth of the average of the order parameter $\langle S_{or} \rangle = \left[\sum_i \langle \boldsymbol{\mu}_i \rangle \cdot \langle \boldsymbol{\mu}_i \rangle \right]^{1/2} / N$, when T^* decreases, suggests that in these two models the dipole orientations freeze to

form a dipolar glass. These results are in disagreement with the mean field theory predictions [160]. The glassy orientational freezing of the disordered quenched XY and DSS models seems to be at variance with the occurrence of fluid ferroelectric phases. This question was further studied in [161, 162] on a model where the dipoles are embedded in mobile soft core spheres whose motion is not influenced by the dipolar interactions. Since the dipole correlations depend on the instantaneous soft core sphere positions, it is possible to value the role of the short range spatial correlations on the onset of the ferroelectric phase. The simulations show that in these dynamically disordered systems the ferroelectric phase exists. They demonstrate the crucial role of the coupling between the dipole orientation and the static or dynamical local spatial order in the existence of stable polarized phases in dipolar fluids.

Susceptibility calculations for DSS with frozen spatial positions (ferrosolid) have been reported by Wang et al. [104] and compared with ferrofluid results. Due to the suppression of cluster formation in the ferrosolid the susceptibilities are lower than in the ferrofluid.

3.5 Dipolar Systems in an External Field

The suspension of dipolar spheres is an ideal, though realistic, model to investigate quantitatively the variation of the susceptibility and magnetization of ferrofluids in the presence of a weak external magnetic field. A detailed analysis of very accurate simulations of a dipolar sphere system (with short range interaction given by a truncated and shifted LJ potential) where the solvent effects are taken into account by coupling the spheres to a Langevin thermostat is presented in [104]. The variation of the susceptibility χ with dipole coupling λ and volume fraction ϕ, typical of those of experimental ferrofluids, is compared with theoretical predictions based on a polynomial expression of χ in terms of the Langevin susceptibility $\chi_L = 4\pi\rho m^2/3kT$ (m magnetic moment) only. Good agreement is obtained for the low dipole moments ($\lambda < 2$) while systematic deviations at larger couplings show that at low and intermediate densities χ depends on λ and ϕ separately. This behaviour is attributed to particle aggregation. Similar discrepancies are observed for the magnetization curves. Additional simulations have been performed for a mixture of two species of dipolar spheres differing by their diameters and dipole moments. The ratio of the dipole moments has been chosen proportional to the ratio of the particle volumes, in accord with the characteristics of magnetite particles. The results described in [163] show that the agreements or differences with theoretical predictions of the magnetization are consistent with those found in the one component fluid. Due to their larger dipole moment, large spheres aggregate more easily in chains than small ones. At a fixed total volume fraction an increase of the volume fraction of the large particles therefore enhances the magnetization and the initial susceptibility.

On the other hand, an increase of the volume fraction of the smaller particles (at a fixed volume fraction of the large particles) hinders aggregation of the large particles and consequently has a significant effect on the structural and magnetic properties of the system [163].

A more realistic particle polydispersity typical of that in a ferrofluid has been considered by Kristóf and Szalai [164]. Their calculations for a DSS model show that, up to moderate fields, the magnetization is generally higher in the polydisperse system than in the corresponding monodisperse system where all particles have dipole moments and sizes equal to the average values.

Comparison of the calculations of initial susceptibility and magnetization performed with different boundary conditions shows that the use of a periodic system with metallic boundary is more appropriate, due to the much weaker dependence on system size, than either a periodic system with vacuum boundary or a finite spherical box [97].

A comprehensive comparison of simulation data for the susceptibility and magnetization of ferrofluid models with theoretical predictions can be found in a recent review article by Huke and Lücke [6].

In an external field sufficiently strongly interacting dipolar particles form chainlike aggregates, aligned with the field direction, whose characteristics depend on density, dipole moment, field strength $\xi = 4\pi mH/(kT)$ and short range interaction. The effect of the field on structural properties is most conspicuous in MC simulations of dipolar hard spheres with centres confined to a plane. In a field parallel to the plane and with low values of λ the chains are short but several chains can loosely bind to form thicker clusters. At larger dipole moments $\mu^* \geq 2.5$ the chains span the simulation cell and aggregate into thicker regularly spaced columnar clusters (stripes) [7, 108, 165]. The average lateral size of the clusters depends on the density increasing from $2.2\,\sigma$ at $\rho^* = 0.2$ to $4.8\,\sigma$ at $\rho^* = 0.4$ to $6.5\,\sigma$ at $\rho^* = 0.6$ [165]. Increasing the field strength favours longer and stiffer chains and produces more densely packed columns. The local structure of the spheres in the columns is that of a hexagonal crystal.

Similar chain ordering has been evidenced in ferrofluid films [166] and inverse ferrofluids confined between narrowly spaced glass plates when a magnetic field is applied parallel to the layer [167–169]. Inverse ferrofluids (or magnetic holes) are nonmagnetic colloidal particles dispersed in a ferrofluid. As the colloidal spheres are much larger (1–100 μm) than the magnetic particles in the ferrofluid (0.01 μm) the ferrofluid can be viewed as a uniform magnetic background. In an external uniform field the spherical colloidal particles of volume V acquire an effective dipole moment $\boldsymbol{m} = -V\bar{\chi}\boldsymbol{H}$ ($\bar{\chi}$ effective susceptibility) collinear with the average magnetic field $\boldsymbol{H}$ inside the ferrofluid.

Three dimensional simulations of strongly interacting DHS in an external field are less extensive but give qualitatively similar results, including aggregation of chains [170]. At large values of the dipole moment high potential

barriers may preclude aggregation into regular columns and the structure appears more gel-like. Columns have been observed to form in thin ferrofluid films (typically a few hundred particle diameters thick) under the influence of a perpendicular magnetic field. They order into a periodic 2D hexagonal lattice with lattice spacing decreasing with field strength [171] and scaling with cell thickness as $L^{1/2}$ [172]. If a strong magnetic field is applied suddenly, a labyrinthine pattern occurs. Similar patterns occur in electrorheological and magnetorheological fluids as will be discussed in more detail in Sect. 5. It can be noted that in thicker films ($\sim 25\,\mu$m, particle diameter $0.089\,\mu$m) the columns develop branched fine structures (tree structure) and the scaling law for the average distance between columns changes to $L^{2/3}$ [172].

The mechanisms through which chains aggregate into bigger clusters are complex, depending on chain length, relative displacement of the chains and lateral distance; they will be discussed in the later Sect. 5 on electrorheological fluids. Chains also form in 3D DSS (in an applied field) but apparently do not cluster [103] possibly due to the softness of the repulsive interaction [173]. Gibbs ensemble MC simulations indicate coexistence between two phases with a critical temperature decreasing with field strength and a critical density practically independent of field strength [103]. Snapshots of configurations at densities on either side of the coexistence curve rule out l-g coexistence and rather indicate the coexistence of two types of chain phases which were, however, not characterized further. Extrapolation to a zero field of the critical parameters gives an upper bound for the critical temperature ($\lambda \sim 14$) and a lower bound for the critical density ($\rho^* \sim 0.032$) for a potential phase transition in DSS. If combined with the upper density bound obtained from the calculations of Leeuwen and Smit [126] ($\rho^* \sim 0.03$) this allows for only a very narrow density range for l-g coexistence in dipolar soft spheres.

As we have seen Stockmayer fluids present a l-g transition as long as the attractive van der Waals interaction dominates the dipolar interaction ($\lambda \leq 25$). In an applied field, at a fixed dipole moment, the critical temperature increases with field strength due to the stronger dipolar interaction (stronger alignment of the dipoles in the field) whereas the density changes only slightly [127]. In the two-phase region liquid droplets, nearly spherical in a zero field, coalesce and become elongated in the direction of the field [127].

Qualitative aspects of the structure induced in DSS by a rotating field ($E_x = E_0 \cos \omega t$, $E_y = E_0 \sin \omega t$) have been investigated by Men et al. [174] and in more detail by Murashov and Patey [175] by MD and Brownian dynamics simulations. The most striking observation is the organization of particles in well-defined layers (perpendicular to the plane of the field) at intermediate values of the angular velocity ω. At smaller or larger values of ω the layers break up and eventually disappear. Snapshots of the orientations of the dipole moments in the layers clearly shows that they are polarized. The formation of layers is favoured at low densities and high dipole moments but also occurs at densities as high as $\rho^* = 0.8$ at sufficiently large dipole moments. The

formation of layers can be qualitatively understood by remarking that if the frequency of rotation is large enough so that the distance between a pair of particles does not change appreciably during one period, the average energy between two dipoles rotating in phase with the field is [175–177]

$$\bar{u}(r) = -\frac{\mu_E^2}{2r^3}(1-3\cos^2\theta_r) \tag{37}$$

where θ_r is the angle between the direction perpendicular to the plane of the rotating field and the interparticle unit vector $\boldsymbol{r}/r$ and it is assumed that the interaction is dominated by the components of the dipole moments $\boldsymbol{\mu}_E$ in the direction of the field and that $|\boldsymbol{\mu}_E|$ is the same for all particles [175]. From Eq. 37 it follows that the average interaction between two particles is attractive when they are in the plane of the field and repulsive when $\boldsymbol{r}$ is normal to the plane which explains qualitatively the layer formation.

Quite interestingly, equimolar mixtures of particles differing only by their moments of inertia can be separated by adjusting the frequency of the applied field. At this frequency only the species with the smaller moment of inertia can follow the field sufficiently well to form layers whereas the other can not [175].

Q2D microsized colloidal suspensions are particularly suited for studying experimentally the formation of ordered structures or melting mechanisms in 2D systems due to the possibility of direct visualization, e.g. through video microscopy.

In a sufficiently strong external electric or magnetic field, perpendicular to the layer, the colloidal particles interact, to a good approximation, through a repulsive long-range $1/r^3$ potential and self-assemble, when confined by walls or magnetic fields, into different planar lattice structures. Triangular (hexagonal) [167, 178–182] but also rectangular, square or oblique lattices (some of which are metastable) have been observed [179].

Melting of the crystal structures has been observed in micrometer-sized colloids confined between glass plates [183] or confined by gravity to a water/air interface [184] and shown to consist of two successive transitions, with an intermediate hexatic phase, as predicted by the KTHNY theory [185–187]. Evidence for hexatic-to-liquid melting has also been given in 2D magnet bubble lattices in magnetic garnet films [188]. These experimental observations are in disagreement with numerical simulations of systems with $1/r^3$ potentials which report first order melting [189, 190].

Melting of 2D clusters of up to 80 dipoles, interacting with a $1/r^3$ potential, confined by a quadratic potential has been determined by means of MC simulations by Belousov and Lozovik [191]. Melting of the smaller clusters (size $N < 37$) differs from that of the larger ones through an additional stage characterized by orientational intershell disordering of pairs of shells.

Orientational fluctuations of the dipole moments in the direction perpendicular to the layer have been investigated in MC simulations by Lomba

et al. [192] for Q2D DHS as a function of density and strength of the perpendicular magnetic field.

Dynamical aspects of Q2D suspensions of spherical colloids and asymmetric colloid mixtures with long-range repulsive forces have been investigated by Brownian dynamics simulations [193–195]. These works emphasize dynamic scaling and the importance of hydrodynamic interactions on self- and tracer diffusion.

3.6 Confined Systems

The confinement in nanoscopic pores can induce important modifications of the thermodynamic properties of simple fluids, for instance, by shifting their temperatures and pressures of liquefaction or freezing. Changes of the location and stability of the ferroelectric fluid phase have been investigated by Klapp and Schoen [196] through MC simulations in an isobaric ensemble for DSS with $\mu^* = 3$, confined in a slit pore constituted of two structureless soft walls. A transverse pressure is applied to the fluid in the direction parallel to the walls. The slit is replicated periodically in the direction perpendicular to the walls and the dipolar interactions summed using the modified Ewald method of Yeh and Berkowitz [87]. A main finding is that, for state conditions where the bulk phase is ferroelectric, the orientational order, which is parallel to the wall, persists down to wall separations as small as 6σ. For a fixed slit width of 7σ it is shown that the confined fluid becomes ordered at pressures significantly lower than in the bulk fluid indicating that confinement promotes ferroelectric ordering. For widths smaller than 6σ, where the dipolar particles are arranged in well-defined layers the existence of a fluid ferroelectric phase could not be clearly identified. One can note that, due to the stratification at these widths, the pressure component normal to the walls has an average value different from the value of the external applied pressure.

A comparison between the properties of a bulk fluid of Stockmayer particles with a permanent dipole moment $\mu^* = 1.8$ and those of the same fluid confined in a spherical cavity with a structureless wall interacting with the dipolar particles by a 9-3 LJ potential has been made for different densities and cavity sizes in [197]. The confinement in the cavity induces strong variations in the density profile near the interface, reduces the value of the dielectric constant and slows down self-diffusion and solvation processes. The authors have extended their simulations to binary mixtures of Stockmayer particles confined by planar surfaces, both charged [198] and uncharged [199], or by metal surfaces [200] focusing on the structural (density and polarization profiles), and dynamic (diffusion) properties of the interfacial particles as a function of surface charges, dipole moments and concentration of the particles. The long-range electrostatic interactions were treated using the slab-adapted Ewald method [87].

Krilov and Laird [201] have investigated the orientational structure of DHS near two parallel discretely polarized hard walls obtained by fixing point dipoles of equal magnitude and equal angle Θ with the wall normal on a triangular lattice. The dipole-dipole interaction was truncated at a distance equal to half of the slit width. A strong dependence on Θ is observed for the orientation of the fluid dipoles near the surface. The first adsorbed interfacial layer is found to be almost crystallized for the highest dipole moment considered ($\mu^* = \sqrt{2}$).

The determination of the properties of the l-g interface of a dipolar fluid has been performed for a Stockmayer system and a system of diatomic particles which, in addition to the point dipole interaction, interact by site-site LJ potentials in [202]. The estimates of the surface tension are shown to be in reasonable agreement with experimental results for 1,1-difluoroethane when state variables are reduced by the critical temperature and density. The preferential orientation of the dipoles is parallel to the interface. This work also contains methodological aspects of the simulation of thin liquid films in equilibrium with their vapour. In particular, a comparison is made between the results obtained for the true (Eq. 25) and slab-adapted Ewald potentials. The agreement between the two numerical determinations of the dipolar energy is quite satisfactory asserting the validity of the use of the 3D Ewald approach for the simulation in a slab geometry.

The adsorption of dipolar fluids in porous solids induces modifications of their bulk properties. Quantitative estimates of these changes on the dielectric properties and spatial correlations of the DHS fluid adsorbed in uncharged or charged disordered or random matrices of immobile hard spheres have been obtained by grand canonical MC simulations [203, 204] for different densities of the adsorbed fluids and porosities of the matrices.

3.7 Dipolar Fluids under Shear Flow

The most detailed simulation study of the orientational ordering of simple dipolar fluids undergoing planar Couette flow at a constant shear rate γ has been presented in a series of papers by McWhirter and Patey [205–208]. In their work the translational motion of the particles is obtained from the so-called SLLOD equations given by [209]

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} &= \boldsymbol{p}_i'/m + y_i\gamma\hat{\boldsymbol{e}}_x \\ \frac{\mathrm{d}\boldsymbol{p}_i'}{\mathrm{d}t} &= \boldsymbol{F}_i - \gamma p_{iy}'\hat{\boldsymbol{e}}_x - \alpha^T[\boldsymbol{p}_i' - m\boldsymbol{u}_s(\boldsymbol{r}_i, t)] \\ I\frac{\mathrm{d}\boldsymbol{\omega}_i^b}{\mathrm{d}t} &= \boldsymbol{T}_i^b - \alpha^R(\boldsymbol{\omega}_i^b - \boldsymbol{\Lambda}_i^b) \end{aligned} \tag{38}$$

where the shear flow is along the x-direction and the velocity gradient along the y-direction. $\hat{\boldsymbol{e}}_x$ is a unit vector along the x-axis, $\boldsymbol{r}_i$ the position of particle i of mass m, $\boldsymbol{F}_i$ the total force acting on it. The total momentum of particle i is $\boldsymbol{p}_i = \boldsymbol{p}'_i + my_i\gamma\hat{\boldsymbol{e}}_x = \boldsymbol{p}''_i + m\boldsymbol{u}_s(\boldsymbol{r}_i, t) + my_i\gamma\hat{\boldsymbol{e}}_x$ so that $\boldsymbol{p}''_i/m$ is the thermal velocity, i.e. the velocity measured with respect to the local streaming velocity $\boldsymbol{u}(\boldsymbol{r}_i, t) = \boldsymbol{u}_s(\boldsymbol{r}_i, t) + y_i\gamma\hat{\boldsymbol{e}}_x$. At low Reynolds numbers and low shear rates, laminar flow is expected to be stable and the streaming velocity is linear, $\boldsymbol{u}(\boldsymbol{r}_i, t) = y_i(t)\gamma\hat{\boldsymbol{e}}_x$. However, at very large shear rates where the Reynolds numbers are large, time-dependent secondary flows $\boldsymbol{u}_s(\boldsymbol{r}_i, t)$ may develop [210]. $\boldsymbol{T}_i^b$, $\boldsymbol{\omega}_i^b$ and $\boldsymbol{\Lambda}_i^b$ are, respectively, torque, angular velocity and streaming angular velocity in the body-fixed frame.

In order to obtain a steady state from Eqs. 38 dissipative heat must be removed from the system. This is achieved by the last (thermostatting) terms of the last two equations in Eqs. 38. In this respect it is essential to observe that accurate values for $\boldsymbol{u}_s$ and $\boldsymbol{\Lambda}_i^b$ are needed. Any deviations from the assumed streaming and angular velocity profiles (biased profiles) will exert unphysical forces and torques which in turn will affect the shear-induced translational and rotational ordering in the system [209, 211, 212]. The values for the multipliers α^T and α^R depend on the particular choice of the thermostat. A common choice, also adopted in the work of McWhirter and Patey, is a Gaussian isokinetic thermostat [209] which insures that the kinetic and rotational energies (calculated from the thermal velocities $\boldsymbol{p}''_i$ and thermal angular velocities $\boldsymbol{\omega}_i^b - \boldsymbol{\Lambda}_i^b$) and therefore the temperature are conserved. Other possible choices are the Hoover–Nosé or Nosé–Hoover-chain thermostats [213–216].

The Eqs. 38 are solved together with Lees–Edwards boundary conditions (LEBCs) [95, 217, 218] where layers of image cells above the central, cubic or parallelepiped, simulation cell slide with uniform velocity γL_y in the positive x direction while those below the central cell slide with velocity $-\gamma L_y$. With these b.c. the new positions $\boldsymbol{r}'_i$ and velocities $\boldsymbol{v}'_i$ of the particles are given in terms of those obtained after integration of the equations of motion by [219]

$$
\begin{aligned}
r'_x &= \begin{cases} (r_x \bmod L_x) + \gamma L_y t & r_y \geq L_y \\ (r_x \bmod L_x) & 0 \leq r_y < L_y \\ (r_x \bmod L_x) - \gamma L_y t & r_y < 0 \end{cases} \\
r'_y &= (r_y \bmod L_y) \\
r'_z &= (r_z \bmod L_z)
\end{aligned}
\tag{39}
$$

$$
\begin{aligned}
v'_x &= \begin{cases} v_x + \gamma L_y & r_y \geq L_y \\ v_x & 0 \leq r_y < L_y \\ v_x - \gamma L_y & r_y < 0 \end{cases} \\
v'_y &= v_y \\
v'_z &= v_z
\end{aligned}
\tag{40}
$$

Implementing the Ewald sums with LEBCs presents no extra difficulty once it is recognized that the system of cubic (or parallelepiped) cells under LEBCs can be mapped onto a monoclinic lattice of parallelepipedic cells. Thus, for a cubic simulation cell, the cell basis matrix would be [218, 220]

$$\boldsymbol{t} = \begin{bmatrix} 1 & \chi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} L \tag{41}$$

where L is the length of the cell side and $\chi = \gamma t$. In [205] the SLLOD equations (38) together with LEBCs and Ewald sums were solved for 256 and 500 Stockmayer particles at reduced temperature $T^* = 1.125$, density $\rho^* = 0.88$ and dipole moment $\mu^* = 1.955$, a thermodynamic state point at which the system is orientationally disordered in the absence of shear. Both conducting and vacuum b.c. were considered corresponding to dielectric constants $\varepsilon' = \infty$ and $\varepsilon' = 1$ of the dielectric medium surrounding the Lees–Edwards lattice, respectively.

Two distinct behaviours are observed depending on the value of the shear rate γ [in units of $(\sigma^2 m/\varepsilon)^{1/2}$]. At low shear rates ($\gamma^{1/2} < 2$), the secondary streaming velocity $\boldsymbol{u}_s$ is negligible, but the streaming angular velocity is not. An orientationally resolved streaming angular velocity seems to be the most suitable choice to prevent external torques to influence ordering in the system [205]. In this regime the nonequilibrium shear viscosity

$$\eta_s(\gamma) = -\frac{\langle P_{yx} \rangle}{\gamma}, \tag{42}$$

where P_{yx} is the off-diagonal element of the pressure tensor, decreases linearly with $\gamma^{1/2}$ extrapolating to the value $\eta_0 = 3.00 \pm 0.05$ ($\varepsilon' = \infty$) and 2.90 ± 0.03 ($\varepsilon' = 1$) at $\gamma = 0$. This value is lower than the corresponding value of the LJ fluid ($\mu^* = 0$) presumably due to a higher barrier to diffusion around a LJ particle than around a dipolar particle [205]. Truncation of the dipole-dipole interaction has a significant effect on η_0, giving a value close to that of the pure LJ system.

The very fact that $\langle P_{yx} \rangle$ is non zero indicates structural distortion of the fluid which manifests by an anisotropic pair distribution function $g(\boldsymbol{r}, \gamma)$ compressed in the xy plane along axes at angles $\phi = 124°$ (ϕ angle between the projection of $\boldsymbol{r}/r$ onto the xy plane and the $\hat{\boldsymbol{e}}_x$ axis) and most elongated along $\phi \sim 34°$. At the value of the dipole moment $\mu^* = 1.955$ considered no strong orientational order is observed and the distortion of the fluid structure results mainly from the LJ part of the potential and is only weakly influenced by the dipolar contribution.

At larger shear rates, the structural properties are found to depend qualitatively on the expression used for the translational thermostat, the value of ε' and the treatment of the dipolar interactions, Ewald sum or truncation [205]. For a *biased* translational thermostat, $\boldsymbol{u}(\boldsymbol{r}, t) = \gamma y_i \hat{\boldsymbol{e}}_x$ ($\boldsymbol{u}_s(\boldsymbol{r}, t) = 0$),

and an angular resolved streaming angular velocity [205], the particles undergo a structural transition, near $\gamma^{1/2} = \gamma_{cr}^{1/2} \sim 2$, forming strings aligned parallel to the flow direction arranged with hexagonal symmetry. Concomitantly the shear viscosity and energy decrease. Above the critical value $\gamma_{cr}^{1/2}$ orientational order of the dipole moments gradually builds up. For conducting b.c. ($\varepsilon' = \infty$), the dipole moments align with ferroelectric ordering in the z direction, the direction where the system appears most compressed. If, instead, vacuum b.c. ($\varepsilon' = 1$) are used two domains form within which dipole moments have opposite orientations (parallel to the z direction). This orientational ordering of the dipole moments changes radically when the dipolar potential is truncated. Now the strings polarize in the flow (x) direction with random orientation of the polarization (up or down) when going from one string to the other.

As at high shear rates neglect of secondary flow may not be justified, an unbiased thermostat obtained as a *truncated* Fourier series representation of a fully 3D $\boldsymbol{u}_s(\boldsymbol{r}_i, t)$, compatible with the Lee–Edwards lattice geometry [211, 221], has been included in Eqs. 38. In this case the string phase disappears for $\gamma > \gamma_{cr}$ due to the appearance of vortices in planes parallel to the yz plane. Orientational ordering is lost, the energy continues to increase for $\gamma > \gamma_{cr}$ and shear thickening rather than shear thinning is observed above γ_{cr} [205]. Still more disturbing, switching to a z-independent form for $\boldsymbol{u}_s(\boldsymbol{r}_i, t)$ restores the string phase for $\gamma > \gamma_{cr}$.

It has been suggested [212] from the simulation results of a LJ system, that the string phase is destroyed if time fluctuations of $\boldsymbol{u}_s(\boldsymbol{r}, t)$ are too rapid at the scale of those of the thermal velocity. Bagchi et al. [212] devised a modified dynamical scheme where, basically, the local streaming velocity $\boldsymbol{u}_s(\boldsymbol{r}, t)$ at $\boldsymbol{r}_i$ in Eqs. 38 is replaced by a dynamical variable $\boldsymbol{u}_{si}$ having corresponding conjugate momentum and mass. By this procedure the time autocorrelation functions of $\boldsymbol{u}_s(\boldsymbol{r}, t)$ and the thermal velocity decay on similar time scales and the string phase is again stabilized. This scheme has not been implemented in the dipolar systems.

Whether the string phase corresponds to a real situation or is an artifact of the simulations due to the use of an inaccurate expression for the secondary flow in regimes where the hydrodynamic stability of planar Couette flow is lost is difficult to ascertain at the present time [205]. Experiments on colloidal suspensions have not provided a clear answer, though at moderate strain rates structural behavior similar to that found in the simulations is observed [222].

At *larger* values of the dipole moments orientational ordering occurs even at low shear rates [207]. If a dipolar soft sphere fluid ($T^* = 1.35$, $\rho^* = 0.80$ and $\mu^* = 2.5$) is subjected to a moderate shear force ($\gamma \sim 0.5$), the dipole moments orient parallel to the xy plane at an angle of approximately 118°, i.e. in the direction where the coordination sphere is most compressed, forming a ferroelectric state. As there is no preference for the net dipole moment to point

up or down the compression axis, it occasionally flips from one state to the other. At some critical value $\gamma \sim 1.7$ orientational order and the shear viscosity decrease. The reason for this decrease is suggested to be related to large fluctuations in the director orientation about the compression axis [207].

Finally, McWhirter and Patey [206] investigated the change in orientational order entailed by planar Couette flow in a DSS system ($T^* = 1.35$, $\rho^* = 0.80$ and $\mu^* = 3.0$) which is ferroelectrically ordered in the unsheared state. Only low shear rates, where $\boldsymbol{u}_s(\boldsymbol{r}, t)$ could be taken as equal to zero, were considered. Contrary to the state with the lower dipole moment where orientational order builds up at low shear rates as a response to spatial structure, here the system looses the orientational order present at $\gamma = 0$ when sheared. The net dipole moment $\boldsymbol{M}(t) = \sum_i \boldsymbol{\mu}_i / |\sum_i \boldsymbol{\mu}_i|$ is shown to rotate continuously, but in a nonuniform way, about the z axis with an average angular velocity roughly equal to the vorticity $-\frac{1}{2}\gamma \hat{\boldsymbol{e}}_z$ and no steady state is obtained. The orientations of $\boldsymbol{M}(t)$, at which the order parameter P_1 drops rapidly, are encountered more frequently at large γ which explains the decrease of $\langle P_1 \rangle$ with increasing shear rate.

Using constraint director dynamics, McWhirter and Patey [206] also determine the shear and twist viscosities describing the coupling between the pressure and shear rate tensors and the Miesowicz viscosities (linear combinations of the former) and show that the latter are qualitatively similar to those of a ferroelectric tetragonal I lattice in accord with the fact that the short-range spatial correlations in the ferroelectric liquid state are similar to those of the tetragonal lattice structure [102].

4 Dipolar Liquid Crystals

Although packing effects alone can drive liquid crystal (LC) phases, as demonstrated by simulations of model hard core LC [223, 224], dipolar interactions, present in a large class of LC materials, can influence the stability of the phases or give rise to phases not observed in nonpolar systems [137]. The majority of simulations has focused on investigating the stability of dipolar LC as a function of position and orientation of an embedded point dipole moment with respect to the molecular axis and its strength. The models most frequently considered are hard spherocylinders (SC), prolate and oblate Gay–Berne (GB) molecules and cut spheres for which the phase diagrams without dipole moment are already known. Reviews of simulation work prior to 1998 can be found in [225] and [226].

The main effect of a central longitudinal dipole moment in SC and GB molecules is to stabilize smectic layering relative to the nematic phase, when compared with the nonpolar case, due to favourable antiferroelectric short range order [225, 227, 228]. As the strength of the dipole increases the nematic

phase eventually becomes unstable indicating the existence of a I-N-Sm triple point [225, 228, 229]. In contrast, the densities of the coexisting isotropic and nematic phases seem only slightly affected by the dipolar interaction for GB particles [227, 229] whereas for hard SC the nematic phase is slightly destabilized relative to the isotropic phase [225]. A plausible explanation is dimerization of the spherocylinders in anti-parallel arrangement thereby reducing the overall aspect ratio driving nematic ordering [225] while for the GB molecules a side-by-side attraction is present already in the absence of a dipole moment [228]. The smectic phase for SC is a S_A phase with random anti-parallel orientation of the dipole moments while for GB particles at elongation $\kappa = 3$ it is S_B (hexagonal bond orientational order in the layers). At elongation $\kappa = 4$ at which, in the nonpolar system, a S_A phase exists between the nematic and S_B phases [230], the dipole moment promotes the S_B phase [228].

In hard SC a central transverse dipole moment also stabilizes the S_A phase relative to the nematic phase but now through head-to-tail association of the predominantly in-plane dipole moments [125]. In the smectic layers the dipole moments form intricated vortex-like chain structures [125] similar to those seen in simulations of Q2D dipolar hard spheres [7, 115].

Shifting the location of the axial dipole moment from a central to a terminal position generally increases the range of stability of the nematic phase [231–234], as compared to the central dipole or nonpolar case, due to the antiparallel dipole pairing at high density yielding aggregates more difficult to accommodate in smectic layers [231]. It also affects the structure of the smectic phase. For a system of polar GB molecules it has been shown that the smectic phase changes, upon increase of the dipole moment, from a well-defined monolayer structure to a strongly interdigitated bilayer structure S_{A_d} due to a strong tendency of anti-parallel association of the dipoles [234]. A large scale simulation of 8000 particles [233] revealed a striped domain structure typical of a $S_{\tilde{A}}$ antiphase smectic [137, 235].

The possibility of formation of a tilted smectic-C phase has been investigated in simulations of GB molecules with two off-centre parallel outwards pointing dipoles having suitable orientations ϕ and $\phi + 180°$ with respect to the molecular axis [236]. At $\phi = 60°$, for a sufficiently low temperature, a tilted S_C phase is identified having hexagonal bond orientational order and tilt ($< 10°$) in the direction of a vertex of the hexagon formed by the nearest neighbours (S_J phase) [236]. A tilted smectic phase is also obtained for $\phi = 0$ but with a tetragonal layer structure (S_T phase). For angles $\geq 60°$ the smectic phase is orthogonal with a hexagonal order (S_B); at $\phi = 90°$ an additional biaxial smectic phase is obtained [236].

None of the prolate dipolar systems mentioned above exhibit (global) ferroelectric ordering. A stable ferroelectric phase has, however, been shown to exist in fluids of oblate ellipsoids of revolution, with a central dipole moment along the axis of revolution, in the range of breadth-to-height ratio

$0 \leq b/a \leq 3$ [148]. One can note that in this range of aspect ratios the orientational order is essentially driven by the dipolar interaction as in their absence the system would be isotropic [223]. For $b/a > 3$, a second transition occurs from a ferroelectric to an antiferroelectric columnar state upon cooling the system [148]. The temperature range over which the ferroelectric liquid is stable narrows with increasing values of b/a and vanishes at $b/a = 0$; for larger aspect ratios the isotropic system transforms directly to the antiferroelectric columnar phase [148]. It should be remarked, however, that ferroelectric behaviour is obtained only in the absence of a demagnetization field ($\varepsilon' = \infty$).

On the other hand, no ferroelectric behaviour is found in disk-like molecules (cut spheres) with point axial dipole moments [237]. In the isotropic and nematic phases the dipole moments do not increase appreciably the short-range correlations already present in the nonpolar cut sphere system and consequently the domain of stability of the isotropic phase is not greatly modified by the dipolar interaction. The columnar phase is marginally stabilized with respect to the nematic phase [237]. In the columnar phase, for axial dipole moments, the hexagonally arranged columns are unpolarized for low dipole moments, and fully polarized for larger moments with antiparallel ordering of neighbouring columns [237].

A biaxial columnar phase could be expected when the dipole moment is transverse to the disk axis, but is observed only in the presence of an additional transverse field [238]. For a transverse dipole moment the columnar structure is found to consist of small groups (~ 3 cut-spheres) with aligned dipole moments organized in an antiparallel fashion along the column [237].

Finally, a tilted columnar phase is induced when the dipole moment makes an angle of 60° with the disk axis. In this phase the tilt angle with the columnar axis is $\sim 20^\circ$; the polarized columns are ordered on a rectangular lattice in an antiferroelectric fashion [237].

Complex smectic phases can result in polar bent-core LC as shown in a MD simulation of a two-site GB model with a transverse dipole moment [239]. For a low bend angle ($\gamma < 20^\circ$) the low temperature phase is a hexagonal tilted S_B phase with the direction of the tilt the same in adjacent layers (synclinic tilt). The tilt angle, which increases with γ, is larger than in the apolar system. The system is overall ferroelectric at $\gamma = 0$ (though the polarization changes in direction during the simulation) and antiferroelectric at $\gamma = 20^\circ$ (adjacent layers have opposite polarizations). For the larger bend angle, $\gamma = 40^\circ$, the low temperature phase is identified as an antiferroelectric biaxial smectic X phase (there is no tilt and no hexagonal order), in contrast to the "TGB-like" biaxial phase found in the apolar model [239].

The combined effect of a dipole moment and a flexible tail, usually present in mesogens, has been addressed in a MC study of hard spherocylinders with a terminal dipole moment and an "ideal" (no volume) flexible endgroup [240]. Adding a flexible tail to a hard core SC acts on the range of stability of the nematic phase in an antagonistic way to the terminal dipole

moment. If the latter enhances the stability of the nematic phase (with respect to the nonpolar case) the former suppresses it, resulting in a narrow range of stability for the nematic phase [240].

5
Electrorheological and Magnetorheological Fluids

Dipolar interactions play an essential role in the so-called electrorheological (ER) and magnetorheological (MR) fluids. ER (MR) fluids are suspensions of dielectric (magnetizable) particles, of typical size 0.5–100 μm, in an insulating (or weakly conducting) nonpolar (nonmagnetic) liquid (e.g. oil) (cf. [2, 3] for a list of electrorheological materials). If the permittivity or permeability mismatch of the dispersed particles and the solvent medium is sufficiently large the electric or magnetic field will induce a dipole moment within the particles proportional to the local field they experience. The structural changes in the fluid (formation of chains and more thick solid structures) entailed by the applied field are responsible for the dramatic (reversible) increase, by several orders of magnitude, of its rheological properties (viscosity, yield stress, shear modulus ...). The mechanical and physical properties will depend strongly on the induced structure and a major aim of simulation of ER and MR fluids is to understand its changes as a function of electric or magnetic field strength and frequency, volume fraction, temperature, and dielectric properties of the dispersed particles and solvent.

5.1
Polarization Model

The most basic model for studying ER fluids, within the context of the polarization mechanism [8, 9], (see [9] for other mechanisms) consists of hard nonconducting dielectric spheres of dielectric constant ε_p and diameter $\sigma = 2a$ (a radius) in a nonconducting continuous liquid of dielectric constant ε_s and viscosity η_s confined between parallel-plate electrodes with separation L_z.

The dipole moment induced in particle i by an external field $\boldsymbol{E}_0$ is

$$\boldsymbol{p}_i = \alpha(\boldsymbol{E}_0 + \boldsymbol{E}_i^{\text{ind}}) = \boldsymbol{p}_0 + \boldsymbol{p}_{i,\text{ind}} \tag{43}$$

where $\alpha = 4\pi\varepsilon_0\varepsilon_s\beta_\varepsilon\, a^3$ is the polarizability, $\beta_\varepsilon = (\varepsilon_p - \varepsilon_s)/(\varepsilon_p + 2\varepsilon_s)$ the dielectric contrast factor ($-0.5 < \beta_\varepsilon < 1$), ε_0 the vacuum permittivity and $\boldsymbol{E}_i^{\text{ind}}$ the electric field created by the multipoles of the neighbouring particles. If the dispersed particles have no net charge and only dipole fields are retained

$$\boldsymbol{E}_i^{\text{ind}} = -\frac{1}{4\pi\varepsilon_0\varepsilon_s}\sum_j \frac{1}{r_{ij}^3}\Big[-3\hat{\boldsymbol{r}}_{ij}\,(\hat{\boldsymbol{r}}_{ij}\cdot\boldsymbol{p}_j) + \boldsymbol{p}_j\Big]\,. \tag{44}$$

The electrical interaction between two particles is

$$u_{ij}^{el} = \frac{1}{4\pi\varepsilon_0\varepsilon_s r_{ij}^3}\Big[- 3(\boldsymbol{p}_i \cdot \hat{\boldsymbol{r}}_{ij})(\boldsymbol{p}_j \cdot \hat{\boldsymbol{r}}_{ij}) + \boldsymbol{p}_i \cdot \boldsymbol{p}_j\Big] \tag{45}$$

and the force [241, 242]

$$\boldsymbol{F}_{ij}^{el} = - \frac{3}{4\pi\varepsilon_0\varepsilon_s r_{ij}^4}\Big[[5(\boldsymbol{p}_i \cdot \hat{\boldsymbol{r}}_{ij})(\boldsymbol{p}_j \cdot \hat{\boldsymbol{r}}_{ij}) - \boldsymbol{p}_i \cdot \boldsymbol{p}_j]\hat{\boldsymbol{r}}_{ij} - (\boldsymbol{p}_i \cdot \hat{\boldsymbol{r}}_{ij})\boldsymbol{p}_j - (\boldsymbol{p}_j \cdot \hat{\boldsymbol{r}}_{ij})\boldsymbol{p}_i\Big] \tag{46}$$

where the $\boldsymbol{p}_i$ must be determined self-consistently by solving Eqs. 43 and 44. In Eqs. 45 and 46, r_{ij} is the distance between the sphere centres and $\hat{\boldsymbol{r}}_{ij}$ the unit vector $\boldsymbol{r}_{ij}/r_{ij}$.

In MR fluids the induced dipole moment is given by an expression similar to Eq. 43

$$\boldsymbol{p}_i = 4\pi\mu_0\beta_\mu\, a^3(\boldsymbol{H}_0 + \boldsymbol{H}_i^{\text{ind}}) = \boldsymbol{p}_0 + \boldsymbol{p}_{i,\text{ind}} \tag{47}$$

where $\boldsymbol{H}_0$ and $\boldsymbol{H}_i^{\text{ind}}$ are the applied and induced magnetic fields at particle i and $\beta_\mu = (\mu_p - \mu_s)/(\mu_p + 2\,\mu_s)$ (μ_p, μ_s relative permeabilities to the vacuum of the particles and solvent, μ_0 permeability of free space).

Eqs. 45 and 46 neglect higher-order multipole moments and many-body contributions beyond the dipole terms and are valid only for large separations of the particles and small values of $\beta_\varepsilon(\beta_\mu)$. Simulation studies often ignore local field effects all together in which case the particles bear a constant dipole moment in the direction of the external field (point dipole approximation).

General methods have been developed to include higher-order multipole moments and many-body effects [243–245] but the computational costs have so far limited their application to simulations of very small system sizes [243].

Higher-order multipole moments enhance the forces between particles at short distances and their neglect is extremely questionable, especially if "fine" effects are looked at, as for instance the ground-state properties of close-packed lattice structures [244, 246–251] or the viscosity. To go beyond the point dipole approximation Klingenberg and co-workers [173, 252] developed an empirical force expression for the interaction between two dielectric spheres in a uniform external field from the numerical solution of Laplace's equation [253]. Recently, Yu and co-workers [254, 255] proposed a computationally efficient (approximate) dipole-induced-dipole model based on a multiple image method which accounts partially for multipolar interactions.

In experimental realizations of ER or MR fluids the system is generally confined between plates (electrodes or glass plates) and these boundary conditions need to be properly taken into account in simulations. In ER fluids where the system is placed between parallel electrodes at constant potential, the transverse electric field must vanish at the electrode boundary. In this

case the image method can be applied to calculate the electric interactions between the fluid particles and the conducting walls. At the point dipole level a dipole of strength p at position (x,y,z) produces an infinite set of image dipoles located at $(x,y,-z)$ and $(x,y,2L_z j \pm z)$, $(j = \pm 1, \pm 2, \ldots)$ and having the same moment [256–258] (origin of coordinates at the lower plate). Each dipole now interacts with its infinite number of images and with the other dipoles in the system and their infinite number of images. The self-energy is given by

$$u_s(z_i) = -\frac{p^2}{\varepsilon_s}\left[\zeta(3)/(4L_z^3) + \sum_{k=-\infty}^{\infty}(2z_i - 2kL_z)^{-3/2}\right] \tag{48}$$

and the interaction of a dipole with another dipole and its infinite number of images by [258, 259]

$$u_i(\rho_{ij}, z_i, z_j) = \frac{p^2}{\varepsilon_s}\sum_{s=1}^{\infty} 4\pi^2 s^2 K_0(s\pi\rho_{ij}/L_z)\cos(s\pi z_i/L_z) \times \cos(s\pi z_j/L_z)/L_z^3 \,. \tag{49}$$

Here $K_0(x)$ is the zeroth-order modified Bessel function of the second type, $\zeta(x)$ the zeta function ($\zeta(3) = 1.2020569\ldots$), $\rho_{ij} = [(x_j - x_i)^2 + (y_j - y_i)^2]^{1/2}$ the lateral distance between two particles and L_z the distance between the two walls.

In the same way as higher multipole moments may affect significantly the short-range particle-particle interaction, they can affect the particle-electrode interaction through multipole images, especially when particles are close to the electrodes [245].

The situation is somewhat different in MR fluids where the confining surfaces can be low permeability materials such as glass plates and a tangential magnetic field is possible. Unlike the ER case there are no image dipoles; the chains remain finite with effective magnetic charges appearing at the chain ends [260, 261]. End charges at a given surface have equal signs; at opposite surfaces, they have opposite signs. The finite length of the chains makes the theoretical calculation of their interaction more difficult [262, 263].

In addition to the electrical interactions, particles interact, at short-range by a repulsive interaction generally taken to be a hard sphere potential. In ER (MR) fluids where the short range repulsions extend over distances much shorter than the size of the particles this is a reasonable approximation. However, in dynamical simulations a discontinuous potential is not convenient and the hard sphere potential is generally replaced by a continuous exponential or inverse power potential with parameters chosen to closely mimic a hard sphere potential. As shown in [173] too soft a potential may lead to unrealistic aggregation of the particles. Potentials of a similar functional form have been used for the wall-particle interaction.

The dynamics of the system can then be obtained from a Langevin-type equation including direct interaction forces $\boldsymbol{F}$ (repulsive, electrical, image and wall-particle forces), hydrodynamic $\boldsymbol{F}^{\mathrm{hyd}}$ and Brownian forces $\boldsymbol{F}^{B}$

$$m\frac{\mathrm{d}^2\boldsymbol{r}_i}{\mathrm{d}t^2} = \boldsymbol{F}_i + \boldsymbol{F}_i^{\mathrm{hyd}} + \boldsymbol{F}_i^{B} \,. \tag{50}$$

For particles suspended in an incompressible Newtonian fluid undergoing linear shear flow the hydrodynamic force can be written [243, 264, 265] (using the notation of [243, 265])

$$\boldsymbol{F}^{\mathrm{hyd}} = -\boldsymbol{R}_{FU}\cdot(\boldsymbol{U}-\boldsymbol{u}^{\infty}) + \boldsymbol{R}_{FE}:\boldsymbol{E}^{\infty} \,. \tag{51}$$

Here $\boldsymbol{F}^{\mathrm{hyd}}$ is a generalized ($6N$ dimensional) force-torque vector, $\boldsymbol{U}-\boldsymbol{u}^{\infty}$ ($6N$ dimensional) is the particle translational-angular velocity relative to the bulk fluid flow evaluated at the particle centre, $\boldsymbol{E}^{\infty}$ (3×3 matrix) is the traceless symmetric rate of the strain tensor (supposed to be constant in space). The resistance matrices $\boldsymbol{R}_{FU}$ ($6N\times 6N$) and $\boldsymbol{R}_{FE}$ ($6N\times 3\times 3$) which depend only on the instantaneous relative particle configurations (position and orientation) relate the force-torque exerted by the suspending fluid on the particles to their motion relative to the fluid and to the imposed shear flow, respectively. Note that in ER (MR) fluids torques can be neglected.

Multibody hydrodynamic interactions have generally been ignored in simulations (for reasons of computational cost) with the notable exception of [243, 264] for a monolayer system involving a small number of particles. Satoh et al. [266, 267] approximate the multibody hydrodynamic forces by assuming additivity of the velocities. This does, however, not guarantee positive definiteness of the mobility matrix (inverse of the resistance matrix), unless a short cutoff radius of the hydrodynamic interactions is used [266, 267].

Hydrodynamic resistance to the motion of the dispersed (spherical) particles from the continuum solvent is generally taken into account through a local friction coefficient

$$\boldsymbol{F}_i^{\mathrm{hyd}} = -3\pi\eta_s\sigma\left(\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} - \boldsymbol{u}_i^{\infty}\right) \tag{52}$$

where $\boldsymbol{u}_i^{\infty}$ is the streaming velocity at $\boldsymbol{r}_i$. For linear shear flow along the x direction and a velocity gradient along the z direction $\boldsymbol{u}_i^{\infty} = \gamma z_i \hat{\boldsymbol{e}}_x$ with γ being the shear rate.

Finally, the effect of collisions of the solvent molecules (thermal forces) is taken into account by a random Brownian force $\boldsymbol{F}_i^{B}(t)$ generally with a white noise distribution (but see [268] for a different choice). Even if Brownian forces are much less than the hydrodynamic and electrostatic forces for sufficiently high fields (> 1 kV/mm) or particle sizes (> 1 μm) they can help the system to overcome getting trapped in local energy minima. Conductivity effects [9, 269] have so far been neglected in simulations.

Multiplying by σ, Eq. 50 can be cast in the dimensionless form

$$\frac{m\sigma^2}{kTt_0}\frac{\mathrm{d}^2\boldsymbol{r}_i^*}{\mathrm{d}t^{*2}} = \frac{F_0\sigma}{kT}\boldsymbol{F}_i^* - \frac{3\pi\sigma^3\eta_s}{kTt_0}\frac{\mathrm{d}\boldsymbol{r}_i^*}{\mathrm{d}t^*} + \frac{3\pi\sigma^3\eta_s}{kT}\gamma z_i^*\hat{\boldsymbol{e}}_x + \boldsymbol{F}_i^{B*} \tag{53}$$

where length, direct force and Brownian force are expressed in units of σ, $F_0 = 3\pi\varepsilon_0\varepsilon_s\sigma^2\beta_\varepsilon^2E_0^2$ (ER) or $3\pi\mu_0\mu_s\sigma^2\beta_\mu^2H_0^2$ (MR) and kT/σ, respectively. In terms of the coupling parameter $\lambda = p_0^2/4\pi\varepsilon_0\varepsilon_s\sigma^3kT$ (ER) or $p_0^2/4\pi\mu_0\mu_s\sigma^3kT$ (MR) one has $F_0 = 3\lambda kT/\sigma$.

If one further chooses the time unit as $t_0 = 3\pi\sigma^3\eta_s/kT$ and parameters η_s, m, σ typical of a realistic ER (MR) fluid, the inertial term in Eq. 53 turns out to be negligible [270–272] and the equation reduces to a first order differential equation

$$\frac{\mathrm{d}\boldsymbol{r}_i^*}{\mathrm{d}t^*} = 3\lambda\boldsymbol{F}_i^* + 8\mathrm{Pe}z_i^*\hat{\boldsymbol{e}}_x + \boldsymbol{F}_i^{B*} \tag{54}$$

which can be solved by Euler's method. Pe is the Peclet number defined as the ratio of hydrodynamic to Brownian (thermal) forces, i.e. $\mathrm{Pe} = 3\pi\eta_s a^3\gamma/kT$ [273, 274]. In the quiescent state (no shear) the second term on the rhs of Eq. 54 is obviously absent.

5.2 Structure of the Quiescent Fluid

The most conspicuous structural change upon application of a sufficiently strong electric or magnetic field in ER or MR fluids is the formation of fibrillous structures parallel to the field. These fiber structures, first observed in ER fluids by Winslow [275] in the 1940s, do not form instantaneously but result from a sequence of nonequilibrium steps with well separated time scales.

The evolution of the system after applying an external field at a given temperature (fixing the parameter λ ratio of electric (magnetic) to thermal forces) will depend on volume fraction ϕ, cell thickness and on whether the field is turned on slowly (low ramping rates), so that thermal motion can allow the system to find its equilibrium state at each field increment, or rapidly in which case the system can easily be locked in local energy minima [276–279]. Although the evolution of ER and MR fluids bears much similarity, especially in the earlier stages, there are qualitative differences due to the different boundary conditions as will be pointed out below.

A succinct description of the structural evolution, from a gas to a columnar solid, as observed in low density ($\phi < 0.3$) ER experiments at equilibrium conditions (slowly increasing field) could be as follows [268, 276, 279, 280].

At low fields ($\lambda < 1$) the electric energies are much smaller than the thermal energies and the system is liquid; the particles are distributed randomly. As the field increases and passes a certain threshold ($\lambda \sim \lambda_{c_1}$) the system forms a nematic liquid. The particles associate into chains of touching

spheres parallel to the field with random positions and lengths. Typical chain formation times in microspheres have been measured [281] to lie in the range of 20–200 ms depending on the material of the spheres and field strength.

At a second critical field ($\lambda \sim \lambda_{c_2}$) the chains grow to reach their nearly maximum length (given by the plate separation) and coalesce into bigger clusters referred to as columns. In ER fluids the columns coarsen with time growing into thicker stable columns with a finite diameter [276]. The structure inside the columns consists of domains of bct crystal which is believed to have the lowest electrostatic energy [244, 247–251, 282]. Column formation is a slower process than chain formation and occurs on a time scale of seconds [276]. Coarsening has been found in light-scattering experiments to vary in time according to a power law behaviour $d \sim t^{\alpha}$ ($\alpha \approx 0.5 \pm 0.1$ [280, 283]) where d is the mean separation between columns, as obtained from the position $2\pi / q_{\max}$ of the main peak in the scattering intensity. This behaviour can be understood semi-quantitatively within the framework of fluctuation-induced coupling between columns [257].

In MR fluids subjected to a slowly increasing magnetic field the evolution from a gas to a columnar solid has been described to consist of four structural regimes [284]. The earlier stages, gas and nematic liquid coexisting with single particles, are quite similar to those in ER fluids [278, 284]. The next two structural regimes, occurring at threshold values of the magnetic field λ_{c_2} and λ_{c_3}, respectively, correspond to columns of loosely bound chains coexisting with single chains and, in the final state, to solid-like densely packed columns with strong local ordering [260, 278, 284]. The coupling constants at the different structural transitions depend slightly on concentration; typical values at $\phi = 0.9$ are 2.42 ($\bar{\lambda}_{c_1}$), 3.79 ($\bar{\lambda}_{c_2}$) and 6.16 ($\bar{\lambda}_{c_3}$), respectively [284].

The spacing between columns approximately satisfies a power law relation with cell thickness $d \sim L_z^{\alpha}$ [260, 285, 286] though α depends somewhat on the range of thicknesses considered to reproduce the experimental results (proposed values are $\alpha \approx 0.37$ [260], 0.50 [286]). Power law dependencies are also observed for the separation between columns with time ($d \sim t^{0.27}$) or magnetic field ($H^{0.22}$) [287] in good agreement with the predictions of a phenomenological model of nucleation-controlled structure evolution [277]. The column spacing depends little on the volume fraction [260, 277] ($d \sim \phi^{1/7}$ [277]) suggesting that the column width increases with ϕ as $\sqrt{\phi}$. Although the columns in MR fluids are likely to solidify into a bct structure [261] they do not coarsen into thick structures as in ER fluids [278] and their shape is not cylindrical [261] but tapered at low density ($\phi \sim 0.03$) or with split ends at higher density ($\phi \sim 0.1$) [284] as a result of the monopole repulsion at the ends of the columns (see below).

When a strong field (e.g. of 1 kV/mm) is applied *suddenly* to an ER fluid, hexagonally ordered sheets, aligned with the field direction, are observed in addition to chains and columns [279]. Sheets are found to grow initially from the electrodes and rapidly (within seconds) form an interconnected

labyrinthine pattern. These metastable sheets slowly (within hours) transform into bct structures which subsequently rearrange into a network of misaligned bct regions [279]. Visualization of the particle arrangement in the sheets reveals zig-zag chains where nearest neighbour pairs make an angle of $\sim 30°$ with the field direction rather than linear chains offset by half a diameter. In this way a layer of particles can form on the electrodes, probably as a result of the strong attraction between the particles near the electrodes and their images. This layer then nucleates the observed structure [279].

Similar complex structures, characterized as folded sheets or bent walls, have been observed in optical microscopy experiments of ferrofluid emulsions at high ramping rates of the magnetic field [288]. When viewed along the field direction they form a labyrinthine pattern whose complexity increases with cell thickness, particle volume fraction and ramping rate [288].

Simulations, even performed with a small number of particles (100–300), based on model Eq. 54 reproduce qualitatively many of the structural features observed in experiment, in particular, the different structural regimes observed in ER experiments, liquid, nematic liquid and bct solid structures [258, 271] thereby giving support to the relevance of the polarization mechanism. The large-scale simulations of Martin and coworkers reveal the importance of defect formation in the aggregation of chains [289] and the "zippering" mode through which lateral chain coalescence can occur [289, 290]; they emphasize the role of thermal fluctuations (Brownian motion), even at high dipole couplings, in obtaining equilibrium structures [268, 291]. When the dipole interaction greatly exceeds kT, Brownian motion is often neglected in the simulations resulting in gel-like (sheets) [270, 274, 289, 292–294] or polycrystalline structures [295] similar to those observed in experiment when a strong field is applied rapidly (see above). The system is kinetically trapped in local energy minima and does not have time to relax during the (often quite short) length of the simulation run. The simulations show that slowly cooling the system by decreasing λ (ramped annealing) instead of a rapid quench is much more effective in developing crystallinity [268]. Differences with experiments have been noted. Thus, for instance, in experiment sheets have not been observed to form from association of chains or to bend into tube or spiral-like structures [279] in contrast to simulation results [289].

An obvious aim of simulations and theoretical studies is to understand the mechanisms by which the structures observed in experiment (or simulations) evolve. A first explanation of how chains can attract each other in ER fluids was provided by Halsey and Toor [257, 297, 298]. These authors remark that once chains have grown sufficiently large to span the electrodes, they behave effectively as infinite chains, due to the image charges. If these chains were perfectly aligned touching spheres with fixed dipole moments in the field direction (no thermal effects) two such (infinite) chains would repel each other if in register and attract each other when offset by one radius in the field direction. However, the electric field created by the perfect chain

is found to decay exponentially in the lateral direction (perpendicular to the field) thereby screening the chain interaction over distances of a few particle diameters.

The reason for aggregation and coarsening is rather to be sought in the disorder created in the chains and columns by either thermal fluctuations [277, 280, 299, 300] or defect formation [289, 291, 300]. According to light-scattering experiments in low density (10^{-5}) ferrofluids chain fluctuations exist even in short chains except for extremely high magnetic fields [301]. The large-scale simulations of Martin et al. [268] have shown that defects are always present in chains and columns even at high temperature. Both chain fluctuations and defects create local variations in the concentration of dipoles which in turn are responsible for the long-range attractive interactions induced between the chains. The latter mechanism is clearly responsible for the coarsening in athermal systems [289], the former is likely to dominate at high temperature where defects can rapidly anneal away or if the field is turned on slowly which limits the formation of defects. The dominant mechanism for chain coalescence can also depend on other factors like concentration, particle roughness or polydispersity [291, 300].

In MR fluids where chains are finite the situation is more complex. Two finite *rigid* chains may attract or repel one another depending on their length (and thus cell thickness and volume fraction), lateral separation and relative shift in the field direction [262, 263, 300]. Even for off registered chains (in contrast to infinite chains) the interaction can be attractive or repulsive because of the repulsion of magnetic charges at the ends of the chains [294]. At short distances the four monopole-monopole interactions associated with the bound charges at the ends of the chains give the dominant contribution to the chain interaction with effect to keep the chain ends apart (tapering). Information on the nature of the interaction between two chains, of length $l = 40\sigma$ free to bend, have been gained from simulation of the attachment time (the faster the latter the stronger the interaction), as a function of shift s of two chains initially separated by a fixed distance (1.7σ) [294]. The strongest attractions are found to occur for $s \leq 10\sigma$ in off registered positions in support of the conical shapes of the column ends. More unexpectedly, for $s > l/2$ registered chains feel the strongest attraction [294].

Beyond some lateral distance (escape distance) two chains, even off registered, repel, this distance increasing with chain length (from 1.2σ at $l = 10\sigma$ to ≈ 2 at $l = 150\sigma$ [294]). This provides an explanation as to why thicker cells (larger chains) at fixed ϕ [260] or higher volume fractions (closer average chain distance) at fixed L_z [260, 302] produce larger aggregates.

From the preceding considerations it is clear that a system of finite chains has a complicated energy landscape with many local maxima and minima in which a rapidly quenched system can be pinned down in line with the complex labyrinthine pattern observed in experiment [288].

The simulations of Mohebi et al. [294] also give insight as to how two chains aggregate when they are attractive. First the ends of the chains bend away from the middle, then the chains make contact from the middle and the two strands on either side attach from the middle to the ends. Zero-temperature calculations confirm this picture [262].

When chains are flexible, as in real systems, thermal fluctuations and defect formation can further increase the escape distance. Direct measurement through optical trap micromanipulation [300] of the lateral interaction of two dipolar chains in a ferrofluid emulsion taking into account their fluctuation behaviour show that long flexible chains are attractive over a rather long-range ($\approx 8\sigma$ at $\bar{\lambda} = 340$) and that in stiffer chains the escape distance is reduced ($\approx 3\sigma$ at $\bar{\lambda} = 610$). These experiments also reveal that at short separations ($4 - 5\sigma$) the force can suddenly shift from attraction to repulsion and back [300].

As already mentioned in this section, in an electric or magnetic field, ER or MR fluids self-assemble into columns with bct crystal structure. This is also the preferred structure in fluids which can be polarized by both an electric and a magnetic field when both fields are parallel or if one field dominates the other [303, 304]. However, when both fields can compete and are applied perpendicularly, the electric field is the first to produce bct columns, then the magnetic field, the system is shown experimentally to develop into a fcc structure [304]. Ground-state calculations of an ideal crystal in the dipole approximation by Tao and Jiang [303] give the hcp structure as the preferred structure, though the energy difference between the hcp and fcc structures is quite small, of the order of the thermal energy at room temperature. Wang et al. [272] have investigated by MD simulations the structural evolution in crossed and compatible electric and magnetic fields in two cases differing by the order the fields are applied. When the two fields are applied simultaneously to an initial random configuration, the particles first form 2D hexagonal layers parallel to the two fields which aggregate into thicker sheet-like structures (at sufficiently high density), then transform into either fcc or hcp close-packed structures (or maybe a mixture of both), depending on the initial configuration and thermal fluctuations [272]. In the second case, closer to the experimental realization of Wen et al. [304], a magnetic field is applied perpendicularly to an initial configuration of cylindrical bct columns. The bct columns rapidly transform into fcc structures which further relax into close-packed structures similar to those in the first case [272].

Similar to the case of permanent DHS [175], MR fluids form layered structures when subjected to a high frequency rotational field [177, 268, 305, 306] but now the structural behaviour results from the interplay of the magnetic and viscous forces. Melle et al. [307] have investigated, by athermal dynamic simulations based on Eq. 54, the low frequency particle dynamics of a dilute strongly interacting 2D system ($\phi = 0.016$) as a function of rotational frequency (which in appropriate reduced units is equal to the Mason num-

ber, the ratio of the hydrodynamic to electrical forces [273, 274], the only relevant parameter in the system). In accord with scattering dichroism experiments [308] a cross-over frequency is found at $\mathrm{Mn} = \mathrm{Pe}/(12\lambda) \sim 1$ (i.e. when hydrodynamic viscous forces overcome the dipolar forces) separating two regimes where the number of aggregate particles and the phase lag between particle structures and field have different behaviour. At frequencies $\mathrm{Mn} < 1$ the chains rotate synchronously with the field but lag behind by a constant angle. The average chain length decreases with increasing rotational frequency scaling approximately as the inverse square of the Mn number in agreement with a theoretical chain model [309]. At $\mathrm{Mn} \sim 1$ the average chain length is approximately two and for $\mathrm{Mn} > 1$ the dimers become unstable and the total number of aggregated particles decreases rapidly.

Recent experimental and simulation work by Martin et al. [310] explored the kind of structures which can be obtained with a triaxial field (e.g. a vertical uniaxial field and a horizontal rotating field) applied to a magnetic suspension. For balanced field amplitudes the dipolar interactions are shown to be dominated by many-body effects leading to surprising structures such as, for instance, stable clusters or particle foam structures.

5.3 Rheological properties

Of major interest are the rheological properties of ER and MR fluids as they condition most of the practical applications. Simulations of ER or MR fluids in a homogeneous shear flow based on Eq. 54 with the hydrodynamic forces given by Eq. 52 have been performed in two [252, 311–313] and three [173, 252, 274, 292, 293, 296, 314–318] dimensions. A constant shear gradient $\gamma = \mathrm{d}v_x/\mathrm{d}z$ in the direction perpendicular to the electrodes can be applied by translating the top electrode in the positive x direction with a speed γL_z, maintaining the bottom electrode fixed. The vorticity direction is along the y axis. In some simulations, to account for the experimental fact that the suspended particles can "stick" to the electrodes, particles near the electrodes are assumed to have the same velocity as the electrode [252, 318, 319]. Only when an electrode friction is present, can a nonvanishing dynamic yield stress be calculated at low shear rates in the polarization model [9].

Although a precise phase diagram may not yet have been obtained from the simulations, different phases have been identified depending on the relative strengths of the Brownian, shear and dipole forces [274, 292, 293]. Convenient dimensionless measures of the relative forces are λ, the Pe and Mn numbers.

In a weak field ($\lambda \leq \lambda_{c_1}$) and low shear rates where the Brownian forces are larger than the dipolar and shear forces, the viscosity is Newtonian (independent of shear rate). When dipolar forces dominate over the Brownian forces the structure of the quiescent state is, as detailed above, bct crystal columns

or gel-like. Under the shear force these structures are pulled in the flow direction, continually rupture and reform, degrade, and, upon an increase of Pe, the particles reorder into flowing layers parallel to the shear gradient-flow velocity (xz) plane with local hexagonal close-packed ordering inside the layers [274, 292, 293, 320]. In the case of bct columns, the close-packed (110) planes in the bct columns are believed to be the slipping layers and the bct $[1\bar{1}0]$ direction the flow direction [274, 293]. At large values of λ, layers may aggregate along the vorticity (y) direction [274, 293].

At the other extreme, when the shear force is the dominant interaction, particles form strings that can slip along the flow direction (shear-string phase) [274, 292, 293]. The projections of the strings on a plane perpendicular to the flow direction form a distorted hexagonal array. The layer and shear-string phases are separated by a region where both phases coexist [293]. In this region dipolar and shear forces are comparable.

When dipolar and Brownian forces compete ($\lambda_{c_1} \leq \lambda \leq \lambda_{c_2}$—this is the region of the nematic liquid phase in the quiescent state) a succession of three phases is apparent upon an increase of Pe [293]: a layered phase at small Pe (Pe$< \lambda/20$), a liquid phase when $\lambda/4 <$ Pe $< \lambda$ and a string phase when the shear force is larger than either the dipolar or Brownian forces (Pe $> \lambda$). A coexistence region exists between the layered and liquid phases.

A periodic layered pattern has been observed at low Mason numbers (Mn $\ll 1$) in MR fluids sheared with an oscillating flow [321, 322]. One can note that a layered stripe phase has also been found in MR fluids in a unidirectional shear flow at Mn ~ 1 where the shearing force dominates the magnetic and Brownian forces [305, 323] as well as in MR suspensions in a rotating field without flow [305].

The steady shear rheological response of ER or MR suspensions is often described by a Bingham-type behaviour with shear stress given by [273]

$$\tau = \eta_{pl}\gamma + \tau_0 \tag{55}$$

where η_{pl}, the plastic velocity, is expected to be accurately given by the high shear rate suspension viscosity η_∞ without field [273]. As noted above, under high shear rates the suspension viscosity is entirely dominated by the hydrodynamic forces and thus insensitive to a variation of the electric field. Eq. 55 incorporates a (dynamic) yield stress [324] defined as the limit of τ as $\gamma \to 0$, as commonly observed in ER experiments at low shear rates ($\gamma < 10^{-3} s^{-1}$).

From Eq. 55 an apparent relative viscosity is given by

$$\frac{\eta}{\eta_s} = \frac{\eta_\infty}{\eta_s} + \frac{\tau_0}{\gamma \eta_s} \tag{56}$$

If one assumes that τ_0 scales like λ η/η_s should depend on the sole Mason number and decrease as $K\mathrm{Mn}^{-1}$. This gives a good representation of experimental data at low and high Mn but not at intermediate values where hydrodynamic and dipolar forces have comparable magnitudes [325].

The simulation results indicate that shear thinning of the viscosity is better represented by a power law [274, 293, 317]

$$\frac{\eta}{\eta_s} = K\mathrm{Mn}^{-\alpha} + \frac{\eta_\infty}{\eta_s} \tag{57}$$

with $\alpha < 1$. However, α is not a universal constant but may depend, for instance, on how the simulations are performed (e.g. at constant λ or constant Pe) [293]. Simulations by Guo et al. [293] in the layer phase predict α to vary from 0.77 at $\lambda = 2$ to 0.98 at $\lambda = 11\,350$, in broad agreement with viscosity measurements by Halsey et al. [326] where α varies from 0.68 ($E = 0.4\,\mathrm{kV/mm}$) to 0.93 ($E = 1\,\mathrm{kV/mm}$). The exponent α also depends on the material of the suspended particles as shown by dielectric measurements [327].

The simulations based on the point dipole model do exhibit a dynamic yield stress [173, 292, 296, 315] and its dependence on the volume fraction and electric field agrees with experiment [173]. However, the magnitude of the yield stress is severely underestimated in comparison with experiment, likely due to the neglect of multipolar and multibody interactions [173, 243].

The influence of particle size on the shear stress has been calculated by Tan et al. [328] for a 2D ER mixture with size ratios 1 : 0.6, 1 : 0.7 and 1 : 0.8 at the constant total area fraction 0.236. The yield stress decreases with concentration of the small particles, reaches a minimum when both species have nearly equal concentration then increases slightly in agreement with experiment [329]. The suspension with the larger particle size has the higher yield stress.

Lukkarinen and Kaski [316] have studied the mechanical properties of ER fluids by simulating perfectly ordered structures (single chains, columns of five chains and bct crystal) under shearing, compression and elongation. They show that ER systems under compressive loading transfer a much larger force from one plane to the other than by shearing.

Compression of a MR fluid after application of the magnetic field produces columns that are thicker and more rigid than the columns grown by natural aggregation (no compression). This change in microstructure can give a yield stress ten times higher than the yield stress without compression [330].

The influence of strain amplitude in an oscillating shear flow (at fixed frequency) on the structural and viscoelastic properties of an ER fluid has been investigated in a MD simulation by Wang et al. [319] and compared with experimental results by Lemaire et al. [331]. With increasing strain amplitude the structure changes from isolated columns to sheet-like structures perpendicular, slanted and parallel to the oscillating direction.

6 Conclusion

Coulomb and dipolar interactions, ubiquitous in nature, have a long range which has to be taken properly into account in systems whose size has been made infinite by periodically replicating the simulation volume in space. Methods based on the solution of Poisson's equation or series summation procedures allow us to derive expressions for the energy of dipolar systems which are rapidly convergent series for most geometries of the simulation volumes, with periodicity in three, two or one spatial directions, commonly used in MC or MD simulations. Use of the Ewald expression for the energy seems the best-adapted way to calculate the energy of systems periodic in all three spatial directions. When associated with numerical methods like PME, SPME, FMM or CMM, simulations of systems of large size, $N > 10\,000$, become feasible. For smaller system sizes the hyperspherical S_3 geometry is an efficient, precise and simple option to simulate dipolar systems.

For 3D systems with periodicity in only two directions the Ewald expression for the energy is computationally expensive and the Lekner expression, with adequate numerical treatment of the singular behaviours, may advantageously be used. However, use of the Ewald expression for systems periodic in 3D with an appropriate surface term also allows for an efficient and accurate evaluation of the energy of systems with a slab geometry.

In general, the precision with which the energy is calculated by the Ewald or Lekner methods can be estimated without ambiguity [56, 81, 88]. Quantitative estimates of the performances obtained for the computation times are more uncertain, depending crucially on the type of computer (processor, cache-memory etc ...), compiler and details of the program coding, though they are certainly correct in a qualitative way e.g. [44, 64, 77, 91].

A large amount of simulation works have been devoted to the study of simple dipolar systems, in particular the DHS and DSS models; the most recent ones focusing on the identification of stable phases have been reviewed. In spite of these numerous contributions quantitative phase diagrams for the two above-mentioned models have not yet been obtained. In particular, finite size effects on the location and stability of the phases have not been assessed. It would seem useful to complement existing data with those obtained with systems involving a few thousand particles which is clearly within reach with present-day computers. An increase of system size will, however, be useful only if accompanied by efficient algorithms, in addition to the standard MC moves, to sample correctly rare events like cluster aggregation, the breaking of chains or rings etc. Due to large metastability of the ferroelectric phase, the location of the transition line remains problematic; an algorithm analogue to the Swendsen–Wang [332] or Wolff [333] algorithms would be necessary to overcome this difficulty but is presently not available for systems with long-range interactions.

Studies like those by McWhirter and Patey on the behaviour of dipolar fluids in a shear flow [205–207] reveal the correlations induced by the long-range of the dipolar interaction since its neglect strongly modifies the characteristics of the orientational order of the dipoles at high shear rates. This work also provides evidence of the sensitivity of the orientational order to deviations from linearity of the streaming velocity.

A comparison of experimental data with those obtained by simulation for the models with interactions described by Eq. 1 remains qualitative taking into account the complexity of the experimental systems; a notable exception are suspensions of PIB-coated iron particles which seem well represented by a DHS model [334].

Present and future simulations will contribute in a decisive way to our theoretical understanding of ER and MR fluids. The importance of the studies undertaken for these systems motivated the review given in Sect. 5 where analogies and differences between the two systems are presented and discussed. Agreement between simulation and experiment is presently only qualitative; progress is clearly necessary, in the simulations, to achieve quantitative agreement. In this respect, including multipolar beyond dipolar interactions and many-body interactions seems indispensable [173, 243, 244].

For reasons of space limitations we have not included in this review simulations of Q2D systems where the dipolar interaction competes with short-range interactions, notable examples being magnetic thin films, layered rare-earth compounds, Langmuir monolayers etc. The orientational ordering in these systems can be quite complex, including spontaneous formation of modulated patterns (stripes, bubbles, intermediate morphologies) resulting from the interplay of long-range repulsive and short-range attractive interactions [82, 335, 336]. Reviews of these simulations, mainly performed for lattice systems, are given in [7, 336]. The present manuscript has been completed in October 2003. More recent work on DHS and ferrofluids can be found in a review paper by Holm and Weis [337].

Acknowledgements We would like to thank A.P. Philipse, G. Bossis, Z. Wang for sending us reprints of their work and G. van Ewijk, K. Butter and G. Koendering for copies of their theses.

References

1. Rosensweig (1985) Ferrohydrodynamics, Cambridge University Press, Cambridge
2. Block H, Kelly JP (1988) J Phys D Appl Phys 21:1661
3. Hao T (2002) Adv Colloid Interface Sci 97:1
4. Hansen JP, McDonald IR (1990) Theory of Simple Liquids, 2nd edn. Academic Press, London
5. Teixeira PIC, Tavares JM, Telo da Gama MM (2000) J Phys: Condens Matter 12:R411
6. Huke B, Lücke M (2004) Rep Prog Phys 67:1731

7. Weis JJ (2003) J Phys Condens Matter 15:S1471
8. Gast AP, Zukoski CF (1989) Adv Colloid Interface Sci 30:153
9. Parthasarathy M, Klingenberg D (1996) Mater Sci Eng R17:57
10. Brush SG, Sahlin HL, Teller E (1966) J Chem Phys 45:2102
11. de Leeuw SW, Perram JW, Smith ER (1980) Proc R Soc A 373:27
12. Lekner J (1991) Physica A 176:485
13. Sperb RP (1994) Mol Sim 13:189
14. Sperb RP (1998) Mol Sim 20:179
15. Sperb RP (1999) Mol Sim 22:199
16. Lekner J (1989) Physica A 157:826
17. Grønbech-Jensen N (1997) Int J Mod Phys C 8:1297
18. de Leeuw SW, Perram JW, Smith ER (1980) Proc R Soc A 373:57
19. Caillol JM, Levesque D (1991) J Chem Phys 94:597
20. Caillol JM (1992) J Chem Phys 96:1455
21. Caillol JM, Levesque D (1992) J Chem Phys 96:1477
22. Caillol JM (1993) J Chem Phys 98:9835
23. Caillol JM (1992) J Chem Phys 96:7039
24. Bystrenko OV (2002) Phys Rev E 65:37702
25. Parry DE (1975) Surf Sci 49:433
26. Parry DE (1976) Surf Sci 54:195
27. Heyes DM, Barber M, Clarke JHR (1977) J Chem Soc Faraday Trans II 73:1485
28. De Leeuw SW, Perram JW (1979) Mol Phys 37:1313
29. Smith ER (1988) Mol Phys 65:1089
30. Rhee YJ, Halley JW, Hautman J, Rahman A (1989) Phys Rev B 40:36
31. Clark AT, Madden TJ, Warren PB (1996) Mol Phys 87:1063
32. Liem SY, Clarke JHR (1997) Mol Phys 92:19
33. Grønbech-Jensen N, Hummer G, Beardmore KM (1997) Mol Phys 92:941
34. Grzybowski A, Gwóźdź E, Bródka A (2000) Phys Rev B 61:6706
35. Grzybowski A, Bródka A (2002) Mol Phys 100:1017
36. Arnold A, Holm C (2002) Chem Phys Lett 354:324
37. Perram JW, Ratner MA (1996) J Chem Phys 104:5174
38. Grzybowski A, Bródka A (2002) Chem Phys Lett 361:329
39. Grzybowski A, Bródka A (2003) Mol Phys 101:1079
40. Lishchuk SV (2002) Mol Phys 100:3789
41. Porto M (2000) J Phys A: Math Gen 33:6211
42. Bródka A (2002) Chem Phys Lett 363:604
43. Kawata M, Mikami M (2001) Chem Phys Lett 340:157
44. Grzybowski A, Bródka A (2002) Mol Phys 100:635
45. Gao GT, Zeng XC, Wang W (1997) J Chem Phys 106:3311
46. de los Santos F, Tasinkevych M, Tavares JM, Teixeira PIC (2003) J Phys Condens Matter 15:S1291
47. Langridge DJ, Hart J F, Campin S (2001) Comp Phys Comm 134:78
48. Boinovich LB, Emelyanenko AM (2003) Adv Colloid Interface Sci 104:93
49. Nymand T, Linse P (2000) J Chem Phys 112:6152
50. Caillol JM, Levesque D, Weis JJ (1989) J Chem Phys 91:5544
51. Aguado A, Madden PA (2003) J Chem Phys 119:7471
52. Perram JW, de Leeuw SW (1981) Physica 109A:237
53. Grønbech-Jensen N (1996) Int J Mod Phys C 7:873
54. Perram JW, Petersen HG, de Leeuw SW (1988) Mol Phys 65:875
55. Kolafa J, Perram JW (1991) Mol Simul 9:351

56. Wang Z, Holm C (2001) J Chem Phys 115:6351
57. Mazars M (2001) J Chem Phys 115:2955
58. Hockney RW, Eastwood JW (1981) Computer Simulations Using Particles, McGraw-Hill, New-York
59. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089
60. York D, Yang W (1994) J Chem Phys 101:3298
61. Luty BA, Tironi IG, van Gunsteren WF (1995) J Chem Phys 103:3014
62. Petersen HG (1995) J Chem Phys 103:3668
63. Sagui C (1999) Annu Rev Biophys Biomol Struct 28:155
64. Toukmaji AY, Sagui C, Board JA, Darden T (2000) J Chem Phys 113:10913
65. Sagui C (2001) J Chem Phys 114:6578
66. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577
67. Darden TA, Toukmaji A, Pedersen LG (1997) J Chem Phys 94:1346
68. Deserno M, Holm C (1998) J Chem Phys 109 7678
69. Greengard L, Rokhlin V (1985) J Comp Phys 60:187
70. Greengard L, Rokhlin V (1987) J Comp Phys 73:325
71. Schmidt KE, Lee MA (1991) J Stat Phys 63:1223
72. Challacombe M, White C, Head-Gordon M (1997) J Chem Phys 107:10131
73. Ding HQ, Karasawa N, Goddard III WA (1992) J Chem Phys 97:4309
74. Christiansen D, Perram JW, Petersen HG (1993) J Comp Phys 107:405
75. Shimada J, Taneko H, Takada T (1993) J Comput Chem 14:867
76. Shimada J, Taneko H, Takada T (1994) J Comput Chem 15:28
77. Lambert CG, Darden TA, Board JA Jr (1996) J Comput Phys 126:274
78. Figueirido F, Levy RM, Zhou Ruhong, Berne BJ (1997) J Chem Phys 106:9835
79. Duan ZH, Krasny R (2000) J Chem Phys 113:3492
80. Kutteh R, Nicholas JB (1995) Comput Phys Commun 86:227
81. Kutteh R, Nicholas JB (1995) Comput Phys Commun 86:236
82. Stoycheva AD, Singer SJ (2002) Phys Rev E 65:36706
83. Hautman J, Klein ML (1992) Mol Phys 75:379
84. Arnold A, Holm C (2002) Comp Phys Com 148:327
85. Shelley JC, Patey GN (1996) Mol Phys 88:385
86. Spohr E (1997) J Chem Phys 107:6342
87. Yeh IC, Berkowitz ML (1999) J Chem Phys 111:3155
88. Arnold A, de Joannis J, Holm C (2002) J Chem Phys 117:2496
89. de Joannis J, Arnold A, Holm C (2002) J Chem Phys 117:2503
90. Kawata M, Mikami M, Nagashima U (2001) J Chem Phys 115:4457
91. Kawata M, Mikami M (2001) Chem Phys Lett 340:165
92. Kawata M, Mikami M, Nagashima U (2002) J Chem Phys 116:3430
93. Mazars M (2002) J Chem Phys 117:3524
94. Predota M, Cummings P, Chialvo AA (2002) Mol Phys 100:2703
95. Allen M P, Tildesley D J (1987) Computer Simulations of Liquids, Clarendon, Oxford
96. Kusalik PG, Mandy ME, Svishchev IM (1994) J Chem Phys 100:7654
97. Wang Z, Holm C, Müller HW (2003) J Chem Phys 119:379
98. Banerjee S, Griffiths RB, Widom M (1998) J Stat Phys 93:109
99. Levesque D, Weis JJ (1994) Phys Rev E 49:5131
100. Weis JJ, Levesque D (1993) Phys Rev Lett 71:2729
101. Wei D, Patey GN (1992) Phys Rev Lett 68:2043
102. Wei D, Patey GN (1992) Phys Rev A 46:7783
103. Stevens MJ, Grest GS (1995) Phys Rev E 51:5962

104. Wang Z, Holm C, Müller HW (2002) Phys Rev E 66:21405
105. Duncan P D, Camp P J (2004) J Chem Phys 121:11322
106. Tavares JM, Weis JJ, Telo da Gama MM (2002) Phys Rev E 65:61201
107. Wen W, Pál KF, Zheng DW, Tu KN (1999) Phys Rev E 59:R4758
108. Satoh A, Chantrell RW, Kamiyama S, Coverdale GN (1996) J Colloid Interface Sci 178:620
109. Davis SW, McCausl W, McGahagan HC, Tanaka CT, Widom M (1999) Phys Rev E 59:2424
110. Frenkel D, Smit B 1996 Understanding Molecular Simulation (San Diego Academic Press)
111. Wu D, Chandler D, Smit B (1992) J Phys Chem 96:4077
112. Camp PJ, Shelley JC, Patey GN (2000) Phys Rev Lett 84:115
113. Camp P J, Patey G N (2002) Phys Rev E 62:5403
114. Tavares JM, Weis JJ, Telo da Gama MM (1999) Phys Rev E 59:4388
115. Weis JJ (1998) Mol Phys 93:361
116. Weis JJ (2002) Mol Phys 100:579
117. Stambach J, Lathrop DP, Ott E, Losert W (2003) Phys Rev E 68:26207
118. Blair DL, Kudrolli A (2003) Phys Rev E 67:21302
119. Butter K, Bomans PH, Frederik PM, Vroege GJ, Philipse AP (2003) J Phys Condens Matter 15:S1451
120. Butter K, Bomans PH, Frederik PM, Vroege GJ, Philipse AP (2003) Nature Materials 2:88
121. Blakemore RP (1975) Science 190:377
122. Philipse AP, Maas D (2002) Langmuir 18:9977
123. Puntes VF, Krishnan KM, Alivisatos AP (2001) Science 291:2115
124. Kachel A, Gburski Z (1997) J Phys Condens Matter 9:10095
125. Gil-Villegas A, McGrother SC, Jackson G (1997) Chem Phys Lett 269:441
126. van Leeuwen ME, Smit B (1993) Phys Rev Lett 71:3991
127. Stevens MJ, Grest GS (1995) Phys Rev E 51:5976
128. Kiyohara K, Gubbins KE, Panagiotopoulos AZ (1997) J Chem Phys 106:338
129. Shelley JC, Patey GN, Levesque D, Weis JJ (1999) Phys Rev E 59:3065
130. McGrother SC, Jackson G (1996) Phys Rev Lett 76:4183
131. Tlusty T, Safran SA (2000) Science 290:1328
132. Zilman AG, Safran SA (2002) Phys Rev E 66:51107
133. Tavares JM, Weis JJ, Telo da Gama MM arXiv:cond-mat/0505043
134. Weis JJ, Levesque D, Zarragoicoechea GJ (1992) Phys Rev Lett 69:913
135. Weis JJ, Levesque D (1993) Phys Rev E 48:3728
136. Gao GT, Zeng XC (2000) Phys Rev E 61:R2188
137. Chandrasekhar S (1992) Liquid Crystals 2nd edition, Cambridge University Press, Cambridge
138. Reske J, Herlach DM, Keuser F, Maier K, Platzek D (1995) Phys Rev Lett 75:737
139. Mamiya H, Nakatani I, Furubayashi T (2000) Phys Rev Lett 84:6106
140. Lagerwall ST (1996) J Phys Condens Matter 8:9143
141. Costa Cabral BJ (2000) J Chem Phys 112:4351
142. Weis JJ (2005) J Chem Phys 123:044503
143. Landau DP, Binder K (2000) A Guide to Monte Carlo Simulations in Statistical Physics Cambridge University Press, Cambridge
144. Nijmeijer MJP, Weis JJ (1995) Phys Rev Lett 78:2887
145. Groh B, Dietrich S (2001) Phys Rev E 63:21203
146. Ballenegger V, Hansen J-P (2004) Mol Phys 102:599

147. Ayton G, Patey GN (1996) Phys Rev Lett 76:239
148. Ayton G, Wei DQ, Patey GN (1997) Phys Rev E 55:447
149. Camp PJ, Patey GN (1999) Phys Rev E 60:4280
150. Blair MJ, Patey GN (1998) Phys Rev E 57:5682
151. Clarke AS, Patey GN (1994) J Chem Phys 100:2213
152. Lavender HB, Iyer KA, Singer SJ (1994) J Chem Phys 101:7856
153. Lu D, Singer SJ (1995) J Chem Phys 103:1913
154. Jund P, Kim SG, Tomanek D, Hetherington J (1995) Phys Rev Lett 74:3049
155. Kund F, Wen W, Pál KF, Tu KN (2001) Phys Rev E 64:61503
156. ten Wolde PR, Frenkel D (1998) J Chem Phys 109:9901
157. ten Wolde PR, Oxtoby DW, Frenkel D (1999) J Chem Phys 111:4762
158. ten Wolde PR, Oxtoby DW, Frenkel D (1998) Phys Rev Lett 81:3695
159. Oh KJ, Gao GT, Zeng XC (1998) J Chem Phys 109:8435
160. Zhang H, Widom M (1995) Phys Rev B 51:8951
161. Ayton G, Gingras MJP, Patey GN (1995) Phys Rev Lett 75:2360
162. Ayton G, Gingras MJP, Patey GN (1997) Phys Rev E 56:562
163. Wang Z, Holm C (2003) Phys Rev E 68:41401
164. Kristóf T, Szalai I (2003) Phys Rev E 68:41109
165. Weis JJ (2005) Mol Phys 103:7
166. Hong CY, Jang IJ, Horng HE, Hsu CJ, Yao YD, Yang HC (1997) J Appl Phys Lett 81:4275
167. Skjeltorp AT (1983) Phys Rev Lett 51:2306
168. Skjeltorp AT (1985) J Appl Phys 57:3285
169. Skjeltorp AT (1995) Physica A 213:30
170. Satoh A, Chantrell RW, Kamiyama S, Coverdale GN (1996) J Colloid Interface Sci 181:422
171. Horng HE, Hong CY, Yeung WB, Yang HC (1998) Appl Optics 37:2674
172. Wang H, Zhu Y, Boyd C, Luo W, Cebers A, Rosensweig RE (1994) Phys Rev Lett 72:1929
173. Klingenberg DJ, van Zwol F, Zukoski CF (1991) J Chem Phys 94:6170
174. Men S, Meunier A, Métayer C, Bossis G (2002) Int J Mod Physics B 16:2357
175. Murashov VV, Patey GN (2000) J Chem Phys 112:9828
176. Halsey TC, Anderson RA, Martin JE (1996) Int J Mod Phys B 10:3019
177. Bossis G, Cebers A (1999) J Magn Magn Mater 201:218
178. Golosovsky M, Saado Y, Davidov D (1999) Appl Phys Lett 75:4168
179. Wen W, Zhang L, Sheng P (2000) Phys Rev Lett 85:5464
180. Pieranski P (1980) Phys Rev Lett 45:569
181. Skjeltorp AT, Helgesen G (1991) Physica A 176:37
182. Helgesen G, Skjeltorp AT (1991) Physica A 180:488
183. Kusner RE, Mann JA, Dahm AJ (1995) Phys Rev B 51:5746
184. Zahn K, Lenke R, Maret G (1999) Phys Rev Lett 82:2721
185. Kosterlitz JM, Thouless DJ (1973) J Phys C Solid State Physics 6:1181
186. Nelson DR, Halperin BI (1979) Phys Rev B 19:2457
187. Young AP (1979) Phys Rev B 19:1855
188. Seshadri R, Westervelt RM (1992) Phys Rev B 46:5142
189. Kalia RK, Vashishta P (1981) J Phys C Solid State Physics 14:L643
190. Bedanov VM, Gadiyak GV, Lozovik Yu E (1982) Phys Lett 92A:400
191. Belousov AI, Lozovik Yu E (2000) Eur Phys J D 8:251
192. Lomba E, Lado F, Weis JJ (2000) Phys Rev E 61:3838
193. Zahn K, Méndez-Alcaraz JM, Maret G (1997) Phys Rev Lett 79:175

194. Pesché R, Kollmann M, Nägele G (2002) J Chem Phys 114:8701
195. Kollmann M, Hund R, Rinn B, Nägele G, Zahn K, König H, Maret G, Klein R, Dhont JKG (2002) Europhys Lett 58:919
196. Klapp SHL, Schoen M (2002) J Chem Phys 117:8050
197. Senapati S, Chandra A (1999) J Chem Phys 111:1223
198. Senapati S, Chandra A (1999) J Chem Phys 110:8129
199. Senapati S, Chandra A (2000) J Chem Phys 113:8817
200. Senapati S, Chandra A (2000) Phys Rev E 62:1017
201. Krilov G, Laird BB (2000) Mol Phys 98:651
202. Mecke M, Fischer J, Winkelmann J (2001) J Chem Phys 114:5842
203. Fernaud MJ, Lomba E, Martin C, Levesque D, Weis JJ (2003) J Chem Phys 119:364
204. Fernaud MJ, Lomba E, Weis JJ, Levesque D (2003) Mol Phys 101:1721
205. McWhirter JL, Patey GN (2002) J Chem Phys 117:2747
206. McWhirter JL, Patey GN (2002) J Chem Phys 117:8551
207. McWhirter JL, Patey GN (2002) J Chem Phys 117:9016
208. McWhirter JL (2003) J Chem Phys 118:2824
209. Evans DJ, Morriss GP (1990) Statistical Mechanics of Nonequilibrium Liquids, Academic, London
210. Evans DJ, Morriss GP (1986) Phys Rev Lett 56:2172
211. Travis KP, Daivis PJ, Evans DJ (1995) J Chem Phys 103:10638
212. Bagchi K, Balasubramanian S, Mundi CJ, Klein ML (1996) J Chem Phys 105:11183
213. Tuckerman ME, Liu Y, Ciccotti G, Martyna GJ (2001) J Chem Phys 115:1678
214. Martyna GJ, Klein ML, Tuckerman ME (1992) J Chem Phys 97:2635
215. Brańka AC (2000) Phys Rev E 61:4769
216. Brańka AC, Wojciechowski KW (2000) Phys Rev E 62:3281
217. Lees AW, Edwards SF (1972) J Phys C 5:1921
218. Wheeler DR, Fuller NG, Rowley RJ (1997) Mol Phys 92:55
219. Wagner AJ, Pagonabarraga I (2002) J Stat Phys 107:521
220. Balasubramanian S, Mundi CJ, Klein ML (1996) J Chem Phys 105:11190
221. Travis KP, Daivis PJ, Evans DJ (1996) J Chem Phys 105:3893
222. Ackerson BJ (1991) Physica A 174A:15
223. Frenkel D (1987) Mol Phys 60:1
224. Frenkel D (1989) Liq Crystals 5:929
225. McGrother SC, Gil-Villegas A, Jackson G (1998) Mol Phys 95:657
226. Gil-Villegas A, Jackson G, McGrother SC (1998) J Mol Liq 76:171
227. Satoh K, Mita S, Kondo S (1996) Liq Crystals 20:757
228. Houssa M, Rull LF, McGrother SC (1998) J Chem Phys 109:9529
229. Houssa M, Oualid A, Rull LF (1998) Mol Phys 94:439
230. Brown JT, Allen MP, Martin del Rio E, de Miguel E (1998) Phys Rev E 57:6685
231. McGrother SC, Gil-Villegas A, Jackson G (1996) J Phys Condens Matter 8:9649
232. Satoh K, Mita S, Kondo S (1996) Chem Phys Lett 255:99
233. Berardi R, Orlandi S, Zannoni C (1996) Chem Phys Lett 261:357
234. Berardi R, Orlandi S, Photinos DJ, Vanakaras AG, Zannoni C (2002) Phys Chem Chem Phys 4:770
235. Levelut AM, Tarento RJ, Hardouin F, Achard MF, Sigaud G (1981) Phys Rev A 24:2180
236. Berardi R, Orlandi S, Zannoni C (2003) Phys Rev E 67:041708
237. Zarragoicoechea GJ, Levesque D, Weis JJ (1993) Mol Phys 78:1475
238. Berardi R, Orlandi S, Zannoni C (2000) Phys Chem Chem Phys 2:2933
239. Johnston SJ, Low RJ, Neal MP (2002) Phys Rev E 66:61702
240. van Duijneveldt JS, Gil-Villegas A, Jackson G, Allen MP (2000) J Chem Phys 92:9092

241. Vesely FJ (1977) J Comp Phys 24:361
242. Wang Z, Lin Z, Tao R (1996) Int J Mod Phys B 10:1153
243. Bonnecaze RT, Brady JF (1992) J Chem Phys 96:2183
244. Clercx HJH, Bossis G (1993) Phys Rev E 48:2721
245. Fu L, Resca L (1996) Phys Rev B 53:2195
246. Tao R, Sun JM (1991) Phys Rev Lett 67:398
247. Tao R, Sun JM (1991) Phys Rev A 44:R6181
248. Friedberg R, Yu YK (1992) Phys Rev B 46:6582
249. Davis LC (1992) Phys Rev A 46:R719
250. Heping Z, Zhengyou L (2000) Physica B 279:174
251. Gross M, Wei C (2000) Phys Rev E 61:2099
252. Klingenberg DJ, van Zwol F, Zukoski CF (1991) J Chem Phys 94:6160
253. Klingenberg DJ, Zukoski CF (1990) Langmuir 6:15
254. Yu KW, Wan JTK (2000) Comput Phys Commun 129:177
255. Siu YL, Wan JTK, Yu KW (2001) Phys Rev E 64:51506
256. Warner M, Hornreich RM (1985) J Phys A Math Gen 18:2325
257. Halsey TC, Toor W (1990) Phys Rev Lett 65:2820
258. Tao R (1993) Phys Rev E 47:423
259. Toor WR, Halsey TH (1992) Phys Rev A 45:8617
260. Liu J, Lawrence EM, Wu A, Ivey ML, Flores GA, Javier K, Bibette J, Richard J (1995) Phys Rev Lett 74:2828
261. Zhou L, Wen W, Sheng P (1998) Phys Rev Lett 81:1509
262. Gross M, Kiskamp S (1997) Phys Rev Lett 79:2566
263. Gross M (1998) Phys Rev E 58:6124
264. Bossis G, Brady F (1984) J Chem Phys 80:5141
265. Banchio AJ, Brady JF (2003) J Chem Phys 118:10323
266. Satoh A, Chantrell RW, Coverdale GN, Kamiyama S (1998) J Colloid Interface Sci 203:233
267. Satoh A, Chantrell RW, Coverdale GN (1999) J Colloid Interface Sci 209:44
268. Martin JE, Anderson RA, Tigges CP (1999) J Chem Phys 110:4854
269. Khusid B, Acrivos A (1995) Phys Rev E 52:1669
270. Hass KC (1993) Phys Rev E 47:3362
271. Tao R, Jiang Q (1994) Phys Rev Lett 73:205
272. Wang Z, Fang H, Lin Zh, Zhou L (2000) Phys Rev E 61:6837
273. Marshall L, Zukoski CF, Goodwin J (1989) J Chem Soc, Faraday Trans 85:2785
274. Melrose JR (1992) Mol Phys 76:635
275. Winslow WM (1949) J Appl Phys 20:1137
276. Wen W, Zheng DW, Tu KN (1999) J Appl Phys 85:530
277. Zhu Y, Haddadian E, Mou T, Gross M, Liu J (1996) Phys Rev E 53:1753
278. Hwang YH, Wu X-l (1994) Phys Rev E 49:3102
279. Dassanayaka U, Fraden S, van Blaaderen A (2000) J Chem Phys 112:3851
280. Martin JE, Odinek J, Halsey TC (1992) Phys Rev Lett 69:1524
281. Wen W, Zheng D W, Tu KN (1998) Phys Rev E 57:4516
282. Chen T, Zitter RN, Tao R (1992) Phys Rev Lett 68:2555
283. Martin JE, Odinek J, Halsey TC, Kamin R (1998) Phys Rev E 57:756
284. Ivey M, Liu J, Zhu Y, Cutillas S (2000) Phys Rev E 63:11403
285. Grasselli Y, Bossis G, Lemaire E (1994) J Phys II (France) 4:253
286. Carletto P, Bossis G, Cebers A (2002) Int J Mod Phys B 16:2279

287. Liu J, Mou T, Zhu Y, Haddadian E, Lin X (1996) In: Bullough WA (ed) Electro-Rheological Fluids, Magneto-Rheological Suspensions and Associated Technology. World Scientific, Singapore
288. Flores GA, Liu J, Mohebi M, Jamasbi N (1999) Phys Rev E 59:751
289. Martin JE, Anderson RA, Tigges CP (1998) J Chem Phys 108:3765
290. Fermigier M, Gast AP (1992) J Colloid Interface Sci 154:522
291. Martin JE, Hill KM, Tigges CP (1999) Phys Rev E 59:5676
292. Melrose JR, Heyes DM (1993) J Chem Phys 98:5873
293. Guo HX, Mai ZH, Tian HH (1996) Phys Rev E 53:3823
294. Mohebi M, Jamasbi N, Liu J (1996) Phys Rev E 54:5407
295. Gulley GL, Tao R (1997) Phys Rev E 56:4328
296. Enomoto Y, Oba K (2002) Physica A 309:15
297. Halsey TC (1992) Science 258:761
298. Toor R (1993) J Colloid Interface Sci 156:335
299. Halsey TC, Toor W (1990) J Stat Phys 61:1257
300. Furst EM, Gast AP (2000) Phys Rev E 62:6916
301. Cutillas S, Liu J (2001) Phys Rev E 64:11506
302. Flores GA, Ivey ML, Liu J, Mohebi M, Jamasbi M (1996) In: Bullough WA (ed) Proceedings of the Fifth International Conference on ER Fluids, MR Suspensions and Associated Technology. World Scientific, Singapore
303. Tao R, Jiang Q (1998) Phys Rev E 57:5761
304. Wen W, Wang N, Ma H, Lin Z, Tam WY, Chan CT, Sheng P (1999) Phys Rev Lett 82:4248
305. Volkova O, Cutillas P, Bossis G, Cebers A, Meunier A (1999) J Magn Magn Mater 202:66
306. Martin JE, Venturini E, Odinek J, Anderson RA (2000) Phys Rev E 61:2818
307. Melle S, Calderón OG, Fuller GG, Rubio MA (2002) J Colloid Interface Sci 247:200
308. Melle S, Fuller GG, Rubio MA (2000) Phys Rev E 61:4111
309. Melle S, Martin JE (2003) J Chem Phys 118:9875
310. Martin JE, Anderson RA, Williamson RL (2003) J Chem Phys 118:1557
311. Klingenberg DJ, van Zwol F, Zukoski CF (1989) J Chem Phys 91:7888
312. Tamura H, Doi M (1992) J Phys Soc Jpn 61:3984
313. Chen TY, Briscoe BJ, Luckham PF (1995) J Chem Soc Faraday Trans 91:1787
314. Melrose JR (1991) Phys Rev A 44:R4789
315. Lapenta G, Maizza G, Palmieri A, Boretto G, Debenedetti M (1999) Phys Rev E 60:4505
316. Lukkarinen A, Kaski K (1998) J Appl Phys 83:1717
317. Martin JE (2000) Phys Rev E 63:11406
318. Gulley GL, Tao R (1993) Phys Rev E 48:2744
319. Wang Z, Lin Z, Tao R (1997) J Phys D:Appl Phys 30:1265
320. Sun JM, Tao R (1996) Phys Rev E 53:3732
321. Cutillas P, Bossis G (1997) Europhys Lett 40:465
322. Cutillas P, Bossis G, Cebers A (1998) Phys Rev E 57:804
323. Volkova O, Cutillas P, Bossis G (1999) Phys Rev Lett 82:233
324. Bonnecaze RT, Brady JF (1992) J Rheol 36:73
325. Klingenberg DJ, Dierking D, Zukoski CF (1991) J Chem Soc Faraday Trans 87:425
326. Halsey TC, Martin JE, Adolf D (1992) Phys Rev Lett 68:1519
327. Lengálová A, Pavlínek V, Sáha P, Stejskal J, Kitano T, Quadrat O (2003) Physica A 321:411
328. Tan ZJ, Zou XW, Zhang WB, Jin ZZ (1999) Phys Rev E 59:3177

329. Wu CW, Conrad H (1998) J Appl Phys 83:3880
330. Tang X, Zhang X, Tao R, Rong Y (2000) J Appl Phys 87:2634
331. Bossis G, Grasselli Y, Lemaire E, Persello J, Petit L (1994) Europhys Lett 25:335
332. Swendsen RH, Wang JS (1987) Phys Rev Lett 58:86
333. Wolff U (1989) Phys Rev Lett 62:361
334. Butter K, Philipse AP, Vroege GJ (2002) J Mag Mag Mat 252:1
335. Seul M, Andelman D (1995) Science 267:476
336. De'Bell K, MacIsaac AB, Whitehead JP (2000) Rev Mod Phys 72:225
337. Holm C, Weis JJ (2005) Curr Opin Colloid Interface Sci (in press)